

AD-A181 830

PREDICTING MILITARY RECRUITER EFFECTIVENESS:  
A LITERATURE REVIEW

Teresa L. Russell and Walter C. Borman  
Personnel Decisions Research Institute

for

Contracting Officer's Representative  
Deirdre J. Knapp

Manpower and Personnel Policy Research Group  
Curtis Gilroy, Chief

Manpower and Personnel Research Laboratory  
Newell K. Eaton, Director

DTIC  
SELECTED  
JUN 22 1987  
S D  
D ✓



U. S. Army

Research Institute for the Behavioral and Social Sciences

April 1987

Approved for public release; distribution unlimited.

87 6 16 017

U. S. ARMY RESEARCH INSTITUTE  
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the  
Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON  
Technical Director

WM. DARRYL HENDERSON  
COL, IN  
Commanding

Research accomplished under contract  
for the Department of the Army

Personnel Decisions Research Institute

Technical review by

Deirdre J. Knapp  
Rebecca M. Pliske

Accession For	
NTIS	CRA&I
DTIC	TAB
Unannounced	
Justification .....	
By .....	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	



This report, as submitted by the contractor, has been cleared for release to Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or other reference services such as the National Technical Information Service (NTIS). The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation.

## UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
ARI Research Note 87-23	AD-A181830	
4. TITLE (and Subtitle) Predicting Military Recruiter Effectiveness: A Literature Review	5. TYPE OF REPORT & PERIOD COVERED Final Report December 85 - November 86	
7. AUTHOR(s) Teresa L. Russell and Walter C. Borman	6. PERFORMING ORG. REPORT NUMBER --	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Personnel Decisions Research Institute 43 Main Street SE, Suite 405 Minneapolis, MN 55414	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2Q263731A792 221C9	
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333-5600	12. REPORT DATE April 1987	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) --	13. NUMBER OF PAGES 65	
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution unlimited.	15. SECURITY CLASS. (of this report) Unclassified	
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) --	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE --	
18. SUPPLEMENTARY NOTES Timothy W. Elig, contracting officer's representative Dierdre J. Knapp, technical monitor This task was performed under a Scientific Services Agreement between the U.S. Army Research Office and Battelle Columbus Laboratories. Task requested (over)		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Army Recruiter Recruiter Selection Recruiter Performance Sales Effectiveness	Environmental Variables Non-cognitive Predictors Cognitive Predictors Criterion Contamination	Training Success Production
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This research note describes research on the identification of factors affecting military recruiter performance. Recruiter selection studies for four branches of the U.S. armed forces are summarized. Within the report, the studies are discussed according to the criteria against which the predictors were validated. The relative merits of various predictors are discussed, and criterion-related issues are described. The note's final section ties together selection research on civilian sales-type jobs, and the research done on selection of military	(OVER)	

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

ARI Research Note 87-23

18. Supplementary Notes (continued)

and funded by the U.S. Army Research Institute.

20. Abstract (continued)

recruiters, in order to further clarify the factors which produce successful recruiting personnel.

UNCLASSIFIED

ii SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

# PREDICTING MILITARY RECRUITER EFFECTIVENESS: A LITERATURE REVIEW

## CONTENTS

---

	Page
INTRODUCTION .....	1
COMMANDING OFFICER NOMINATIONS AND SUPERVISOR RATINGS AS MEASURES OF RECRUITER PERFORMANCE .....	4
Wollack & Kipnis (1960) .....	4
Krug (1972) .....	6
Abrahams, Neuman, & Rimland (1973) .....	7
Brown, Wood, & Harris (1975) .....	8
Graf & Bower (1976) .....	9
Borman, Hough, & Dunnette (1976) .....	10
Borman, Toquam, & Rosse (1979) .....	10
Borman, Rosse, & Abrahams (1980) .....	12
Borman, Rosse, & Rose (1983) .....	15
PRODUCTIVITY AS A MEASURE OF RECRUITER EFFECTIVENESS .....	18
Production as Criteria for Validating Predictors .....	18
Brogden & Taylor (1949) .....	18
Larriva (1975) .....	19
Atwater, Abrahams, & Trent (1986) .....	21
Environmental Contaminants of Production .....	22
Bennett & Haber (1973) .....	22
Brown, Wood, & Harris (1975) .....	24
Arima (1977) .....	25
Borman, Rosse, & Toquam (1982) .....	26
PREDICTING RECRUITER SUCCESS IN TRAINING .....	28
Massey & Mullins (1966) .....	28
Borman (1982) .....	30
Borman, Eaton, Bryan, & Rosse (1983) .....	34
MEASURING AND PREDICTING PERFORMANCE IN MILITARY RECRUITER AND CIVILIAN SALES JOBS: AN INTEGRATION .....	36

CONTENTS (continued)

	Page
The Performance Domain .....	37
Measuring Performance and Effectiveness .....	38
Objective Measures .....	40
Subjective Measures .....	41
Personal Attributes Important for Effective Performance .....	43
Biographical Variables .....	43
Personality Variables .....	44
Vocational Interest Variables .....	45
Skill Level Variables .....	46
Aptitude Variables .....	46
Summary: Predictor-Criterion Relationships .....	47
REFERENCES .....	48
APPENDIX A. SUMMARY OF THE DEVELOPMENT OF THE SPECIAL ASSIGNMENT BATTERY (SAB) .....	55

LIST OF TABLES

Table 1. Predictor-Criterion Relationships Investigated in Military Recruiter Research .....	3
2. SAB Composites/Keys (Pooled) Validity Estimates ...	13
3. Factors Identified for Each Performance Category ..	14
4. Comparison of Rural and Urban Recruiter Samples After Application of Regression Equations .....	20
5. Means, Standard Deviations, and Sample Sizes for the Predictive Recruiter Study.....	23
6. Correlations Between Assessment Measures and Training Criteria .....	33
7. Recruiter and Civilian Sales Job Analyses.....	39

## PREDICTING MILITARY RECRUITER EFFECTIVENESS: A LITERATURE REVIEW

### EXECUTIVE SUMMARY

---

#### Requirement:

The U.S. Army has a continuing need to recruit sufficient numbers of qualified men and women into service. Recruiters responsible for contacting prospective recruits and selling them on an Army career are clearly key personnel for enlisting qualified persons. The selection of recruiters with strong potential for success on this job is, therefore, important. The purposes of the present report are three-fold: (1) to provide information on the individual differences variables likely to be successful predictors of military recruiter performance; (2) to discuss issues related to measuring recruiter performance; and (3) to identify research on civilian sales selection that is relevant to recruiter selection research.

#### Procedure:

Past research on military recruiter selection and on recruiter performance measurement is reviewed. The results are compared to similar research for sales jobs in the civilian sector.

#### Findings:

Three major findings arose from analysis of recruiter and civilian sales selection research.

First, non-cognitive measures have successfully predicted military recruiter performance. More specifically, past research suggests that vocational interest and personality variables (such as dominance, self-confidence, and spontaneity) are significantly associated with military recruiter performance. Cognitive variables such as verbal ability and general aptitude have been employed less frequently in recruiter selection research and appear to have little validity for predicting military recruiter success.

Second, military recruiter performance has been evaluated using supervisor, peer, and self ratings, production data, and measures of training success. Each type of criterion has advantages and disadvantages.

Although recruiter productivity (i.e., the gross number of recruits over a specified time frame) is an obvious performance criterion, production data may be influenced or contaminated by environmental factors outside the recruiter's control such as the unemployment rate in the territory. Additionally, production is an index that may not tap some aspects of performance. In contrast, peer, self, and supervisory ratings are probably not restricted by environmental factors and can provide multidimensional evaluations of performance. These features offer a distinct advantage in selection research. On the other hand, the quality of such subjective ratings is influenced by the source of the rating and various rating errors and biases. Training criteria are different conceptually, and often methodologically, from measures of success on-the-job. That is, training performance may be related to behaviors such as studying that may not be highly related to on-the-job performance. In turn, variables that predict job performance may differ from those that are related to training success.

Finally, although military recruiting is essentially a sales job, the type of "product" sold distinguishes this job from sales jobs in the civilian sector. In particular, recruiters are selling a career or life-style rather than material goods or services. Research in the civilian sector suggests that differences in the type of product sold moderates predictor-criterion relationships. Even though "product" type differences should produce different results for military recruiter and civilian sales jobs, there are two major consistencies between factors that contribute to military recruiter and civilian sales success. First, personality variables such as dominance have yielded reasonable validities in both arenas. Second, aptitude and verbal ability measures have shown little merit for predicting either civilian sales or military recruiter performance. Additionally, research from the civilian sector and, to a limited extent, military recruiter selection research suggests that skill level variables such as assessment center scores may be useful predictors of recruiter job performance.

#### Utilization of Findings:

The present report provides a review of all previous military recruiter selection research. Moreover, the information provided in this report can be used for at least three purposes: (1) to help identify variables likely to be successful predictors in future selection research; (2) to provide guidance in criterion development work related to the Army recruiter job; and (3) to aid in identifying factors likely to influence or contaminate production data. In sum, researchers contemplating future work in this area should first carefully review the issues in military recruiter criterion measurement and the conclusions about predictors of military recruiter performance presented in this report.

## PREDICTING MILITARY RECRUITER EFFECTIVENESS: A LITERATURE REVIEW

### Introduction

Over the past 20 years, the U.S. Armed Forces' all-volunteer concept has placed considerable pressure on recruiters to attract sufficient numbers of qualified recruits. Recruiters can no longer depend on selective service to help them fill personnel quotas, and must instead sell military service as an alternative to civilian job opportunities. Additionally, military recruiters must attract to the service individuals who either already possess technical skills or have good potential for developing skills necessary to do technical jobs. The percentage of military personnel who must be technically proficient is growing and will continue to grow in the future as the services continue to modernize their equipment and jobs. Thus, effective recruitment requires enlisting sufficient numbers of high quality recruits.

An important question for the military services, therefore, is how to maintain an effective recruiting force. Sales and other kinds of training is clearly one way to enhance and upgrade the performance of recruiters. Another vehicle is selection of individuals who possess the requisite skills, abilities, and personal characteristics to become successful military recruiters. Regarding selection, the Army, Navy, Marine Corps, and Air Force have conducted a number of studies to identify tests and other kinds of measures that might be used to predict success in military recruiting. The present paper is a review of this work.

Research on military recruiter selection has typically involved administering to a sample of recruiters one or more predictor measures, such as ability tests, personality scales, or vocational interest inventories and also assessing in some manner the performance of these same recruiters. Relationships between the predictor and performance criterion measures have then been used to evaluate the validity of the predictors. In most cases, a concurrent validation strategy has been employed in which predictor measures are administered to recruiters who are already on the job and at or about the same time the criterion measure(s) is taken.

As mentioned, selection research on military recruiters has involved several different kinds of predictor measures. Also, different types of performance criteria have been employed in this work. Performance ratings or nominations made by supervisors or peers, production data (i.e., how many recruits per unit time are brought into the service), and training performance criteria have all been used to assess the validity of predictor measures. The present report is organized according to these three different kinds of criteria. Each chapter reviews and evaluates the validity of the various kinds of predictors against one of these criterion types.

The reason for this structure is that validity of a predictor depends to some extent on the particular criterion employed. The three types of criteria are sufficiently different in nature that a predictor could relate well to performance utilizing one of the criteria but not as well with another criterion. Further, each of these three kinds of measures has

merit as a criterion and is therefore worth trying to predict. Ratings by superiors, for example, reflect informed opinions about recruiter performance based on good opportunity to observe recruiter performance-related behavior over a period of time. Production, number of recruits brought into the service per unit time, is a "bottom-line" index of a recruiter's effectiveness, and success in training is clearly important as a prerequisite for performing effectively later on the job.

To obtain information on military recruiter selection research, we contacted representatives of the four branches of the Armed Services and requested copies of technical reports. Studies are discussed chronologically within each chapter, and for each study a full summary is provided including a description of the predictor measures and criteria utilized and an evaluation of the results and conclusions.

The purpose of this review is to explicate the merits of different types of predictor measures (e.g., cognitive, non-cognitive) for predicting various aspects of performance in recruiting. Studies reviewed in this paper are listed according to the types of predictors and criteria employed in Table 1, and these predictor-criterion relationships are summarized at the end of each chapter. The final chapter of this paper incorporates civilian sales selection research to provide a broader perspective of factors that influence effectiveness in the military recruiter job.

Table 1  
Predictor-Criterion Relationships Investigated in Military Recruiter Research

Predictor Type	Criterion Type						Other Turnover/ Job Survival	
	Performance			Production				
	Commanding Officer Nominations	Peer, Self, or Supervisor Ratings	Unadjusted	Adjusted	Pass/Fail	Test Scores		
Non-Cognitive	Wollack & Kipnis (1960)	Wollack & Kipnis (1960)	Brogden & Taylor (1949)	Larriva (1975)	Massey & Mullins (1966)	Borman (1982)*	Brogden & Taylor (1969)	
	Krug (1972)	Massey & Mullins (1966)	Bennett & Haber (1973)	Borman, Toquam, & Rosse (1979)	Borman, Eaton, Bryan, & Rosse (1983)**	Borman, Eaton, Bryan, & Rosse (1983)**	Atwater, Abrahams, & Trent (1986)	
	Abrahams, Newman, & Rimland (1973)	Graff & Bower (1976)	Borman, Rosse, & Brown, Wood, & Ross (1980)	Borman, Toquam, & Ross (1979)	Borman, Eaton, Bryan, & Rosse (1983)**	Borman, Eaton, Bryan, & Rosse (1983)**		
	Brown, Wood, & Harris (1975)	Borman, Rosse, & Atwater, Abrahams, Ross (1980)	Borman, Rosse, & Atwater, Abrahams, & Trent (1986)	Borman, Rosse, & Rose (1983)	Borman, Eaton, Bryan, & Rosse (1983)**	Borman, Eaton, Bryan, & Rosse (1983)**		
		Borman, Rosse, & Abrahams (1980)	Borman, Rosse, & Atwater, Abrahams, & Trent (1986)					
Cognitive	Wollack & Kipnis (1960)	Wollack & Kipnis (1960)	Massey & Mullins (1966)	Massey & Mullins (1966)	Borman (1982)*	Borman (1982)*	Borman (1982)*	
	Brown, Wood, & Harris (1975)				Borman, Eaton, Bryan, & Rosse (1983)**	Borman, Eaton, Bryan, & Rosse (1983)**	Borman, Eaton, Bryan, & Rosse (1983)**	

\*A paper-and-pencil non-cognitive inventory and assessment center exercises were both employed in this study as predictors.

\*\*\*Assessment center exercises were employed as predictors in this study.

## Commanding Officer Nominations and Supervisor Ratings as Measures of Recruiter Performance

In several studies, researchers have obtained nominations of effective and ineffective recruiters from commanding officers and used these nominations as criteria for assessing the validity of predictors. Likewise, supervisors' ratings of recruiter effectiveness have been gathered and applied as criteria for predictor validation. This section describes studies employing commanding officer nominations and/or supervisor ratings as criteria for success in recruiting.

Wollack and Kipnis (1960). Commanding officer nominations and supervisor ratings served as criteria in a validation and cross-validation study conducted by Wollack and Kipnis (1960). These researchers developed a Navy recruiter selection battery composed of 13 measures designed to reflect interests in recruiting activities, fluency of expression, general aptitude, and knowledge of the Navy. The predictor battery included:

1. Fluency of Expression: Four timed inventories designed to measure a subject's verbal abilities. (These timed inventories were administered to the cross-validation group only.)
  - a. Thing Listing Test - requires the subject to write names of as many things as possible.
  - b. First Letters Test - requires the subject to write as many words as he/she can think of beginning with the letter "b".
  - c. Four Word Combination Test - requires the examinee to compose four word sentences using the four letters presented in each item as the initial letters of the words.
  - d. Inventive Opposite Test - a word is presented along with the first letter of one of its synonyms and the first letter of an antonym. The subject is required to fill in the synonym and antonym for each word in the series.
2. Knowledge of and Enthusiasm for the Navy
  - a. Navy Knowledge - a 60-item inventory designed to measure a subject's knowledge of Navy history and traditions.
  - b. Career Preference Scale - 30 items intended to measure the subject's attitude toward a Navy career.
  - c. Career Motivation Survey - a 39-item attitude scale, measuring attitudes toward various aspects of Navy life.
3. Vocational and Sports Interests
  - a. Kuder Preference Record - a published inventory designed to measure a subject's interest in ten broad occupational areas.

- b. Sports Inventory - a 50-item inventory dealing with rules and plays of various sports, designed to measure a subject's interest in masculine activities.

4. Basic Test Battery: Measures aptitude in four areas:

- a. General Intelligence
- b. Arithmetic Reasoning
- c. Mechanical Comprehension
- d. Clerical Skills

The validation sample consisted of 410 active recruiters representing 40 recruiting stations.<sup>1</sup> This sample was formed by contacting the commanding officers from the 40 recruiting stations and asking each to nominate the most effective and least effective recruiters from their respective stations. These nominations were employed as the criterion measure against which the predictors were validated (most effective recruiter group  $N = 205$ , least effective recruiter group  $N = 205$ ).

Item analyses were conducted for The Navy Knowledge Test, Career Preference Scale, Career Motivation Survey, Kuder Preference Record, and Sports Inventory using the effective and ineffective nominations as the criterion. Items that discriminated between effective and ineffective recruiters beyond the .20 level of confidence were retained for cross-validation. Seven items from the Navy Knowledge Test discriminated beyond the .20 level, as did 15 items from the Career Preference Scale, 24 items from the Career Motivation Survey, 13 items from the Sports Inventory, and four scales from the Kuder.

Navy service rating classification (duty prior to serving as a recruiter) was also analyzed, and significant differences were found to exist between the effective and ineffective groups. The effective recruiter group contained larger proportions of men in Deck ( $p < .01$ ) and Aviation specialties ( $p < .01$ ). The ineffective group contained significantly greater proportions of men in Engineering and Hull specialties ( $p < .01$ ) and in Construction jobs ( $p < .02$ ).

The sample used to cross-validate the predictor battery consisted of 260 students attending a six-week recruiter course. The predictor battery was administered to the cross-validation group while they were attending the course. Approximately one year after the initial testing, when all students had been assigned to recruiting duty, ratings were collected from the two supervisors most familiar with each recruiter. The evaluation

---

<sup>1</sup>The Navy term "station" is equivalent to the Army recruiting "station" level. Navy districts are approximately equivalent to the battalion level, and the Navy term "zone" is equivalent to the Army term "company."

forms contained four 14-point scales: Technical Competence, Willingness to Work, Military Manner, and Adaptability. The forms also required the supervisors to answer the following two questions using 5-point scales:

1. If you had been given the chance would you have recommended the man for recruiting duty?
2. Is this man effective in recruiting personnel?

The six ratings on the evaluation scales were treated as six different criterion measures. Recruiter scores on each dimension were computed by adding the two supervisory ratings together. When only a single supervisory rating was available, the dimension scores were simply doubled. A total of 222 recruiters were rated by two supervisors and interrater reliabilities ranged from .56-.66, indicating an acceptable level of rater agreement.

Product-moment correlations between the predictors selected previously and the criterion measures were computed using the cross-validation sample of 260. Three of the variables significantly predicted the overall evaluation of recruiting performance: The Kuder Persuasive Scale ( $r = .24$ ,  $p < .01$ ), the Kuder Scientific Scale ( $r = -.17$ ,  $p < .01$ ), and the Career Motivation Survey item analysis key ( $r = .13$ ,  $p < .05$ ). None of the other predictors correlated significantly with any of the criteria with the exception of one fluency of expression measure, the First Letters Test, which yielded a correlation of  $-.18$  ( $p < .01$ ) with Military Manner. A multiple correlation was computed using the Persuasive, Scientific, and Career Motivation survey predictors against the criterion measure of overall recruiter effectiveness. The result ( $r = .26$ ) was not significantly greater than the correlation between that performance dimension and the Persuasive Scale alone. Thus, results of the cross-validation analysis suggest that an effective Navy recruiter has persuasive interests, is not very interested in scientific pursuits, and believes in the value of a Navy career.

However, overall, results here were disappointing. Very few of the wide variety of cognitive and non-cognitive variables correlated significantly with performance. Results regarding the cognitive predictors were particularly disappointing in that only one out of 24 predictor-criterion correlations for the verbal fluency tests was significant, and none of the general aptitude tests showed evidence of validity.

Krug (1972). In another study (Krug, 1972), both biographical and personality items were included in a battery designed to predict Navy recruiter performance. Krug administered to Navy enlisted and officer recruiters the 16PF personality inventory, a 25-item inventory supplement (including a 7-item motivational distortion scale and 18 items designed to measure strength of motivation toward a career as a Naval recruiter), and seven items tapping demographic information. Commanding officers from each of 41 Naval recruiting stations across the continental United States were asked to nominate five enlisted recruiters and one officer recruiter from among the top 50 percent performers and five enlisted recruiters and one officer recruiter from among the bottom 50 percent performers of recruiters presently on duty. These nominations of effective and ineffective recruiters served as the performance criterion against which the 16PF,

supplemental, and demographic variables were validated (N = 383 enlisted recruiters; N = 74 officer recruiters).

A stepwise multiple regression of the predictor variables on the criterion resulted in a multiple correlation coefficient of .40 for the enlisted sample, significant beyond the .01 level. When the same equation was applied to the officer recruiter cross-validation sample, a multiple correlation coefficient of .25 was obtained ( $p < .05$ ). According to these results, the effective Navy recruiter is typically married, has more years of formal education, tends to be warm and outgoing, dominant, aggressive, self-assured, and politically conservative.

Although the results from the cross-validation sample supported the validity of this instrument, it is important to take a closer look at the regression equation. The most influential variable, Marital Status, was assigned a high positive weight in the prediction equation. Yet, only 13 of the 383 enlisted recruiters in the sample were unmarried. In a selection situation, a low base rate of unmarried recruiters would make that variable less useful than a variable with a more balanced base rate.

Abrahams, Neuman, and Rimland (1973). Another study focusing on the utility of non-cognitive predictors for selecting Navy recruiters was conducted by Abrahams, Neuman, and Rimland (1973). These authors employed the Strong Vocational Interest Blank (SVIB) and an officer nomination criterion to develop an empirically keyed Recruiter Interest Scale (RIS). The commanding officers in each of 42 recruiting districts were requested to identify the five least effective and five most effective recruiters, and SVIBs were mailed and administered to each of these 420 recruiters.

Complete information was obtained from 356 recruiters representing 36 of the recruiting districts. One half of the sample was used to develop the RIS while the other half was used as the cross-validation group. Each of the 399 SVIB items was weighted according to the proportion of most and least effective recruiters endorsing the response. If a response was endorsed more often by the most effective recruiters it was weighted +1. Conversely, if a response was endorsed more often by the least effective recruiters it was weighted -1. The 115 items with the largest endorsement differences were included in the SVIB-Recruiter Interest Scale (RIS-1).

When the "holdout" group was used to cross-validate the SVIB-RIS-1, the scores for each recruiter were tabulated and arranged from high to low (high scores indicating potentially effective recruiter performance and low scores indicating potentially ineffective recruiter performance). The cross-validation sample was then divided into quartiles according to RIS scores. The top quartile (highest RIS scores) contained three times as many effective recruiters as did the bottom quartile; conversely, the bottom quartile contained three times as many ineffective recruiters as did the top RIS group. Based on these results, the authors recommended using the SVIB-RIS-1 as a selection device to exclude low scoring individuals from recruiter service.

The authors recommended using supplemental inventories (e.g., biographical items found effective in predicting sales performance) to increase validity. Abrahams et al. also stressed that a better criterion of recruiter effectiveness must be developed and that specific elements of

recruiter performance (e.g., prospecting, closing the sale) should be taken into account as criteria in future validity studies.

Brown, Wood, and Harris (1975). Brown, Wood, and Harris developed an Army recruiter selection battery composed of inventories expected to assess personal characteristics thought to be important for success as a recruiter. The selection battery consisted of 15 paper-and-pencil inventories and one verbal performance test. Below is a list of the measures included in the battery.

1. Verbal Fluency--Recruiters were asked to simulate a presentation of the benefits of Army life to a prospective enlistee. Each presentation was recorded and scored by computing the ratio of the number of words spoken during the first two minutes of the presentation to the number of "ahs" spoken during the same length of time.
2. Sociability Measures--Four inventories were used to measure a recruiter's sociability and affiliative tendency.
3. Achievement Motivation--Three inventories were employed to tap the tendency to work hard to achieve self-appointed goals.
4. Empathy Measures--Four instruments were used to measure the ability to understand the point of view of others and the drive to win or complete a sale.
5. Rejection Tolerance Measure--One paper-and-pencil inventory was employed to measure a recruiter's tolerance to rejection, rebuffs, and insults.
6. Responsibility and Maturity Measures--Three instruments were used to tap information about a recruiter's ability to manage his personal and financial duties.

To supplement the predictor battery, aptitude test scores and biographical information were gathered from Army personnel files.

As a criterion performance index, Brown et al. devised a Composite Supervisory Rating procedure to select two extreme groups of recruiters--the highly successful and the highly unsuccessful.<sup>2</sup> The highest ranking District Recruiting Command (DRC) and the lowest ranking DRC (based on objectives achieved) were selected from each of the five Regional Recruiting Commands. Five supervisors from each DRC were then asked to nominate the ten best and ten poorest recruiters within their DRC. The

---

<sup>2</sup>These authors also developed an adjusted production criterion (as described in the next section); however, when the validation study began, the development of the production criterion was not completed. Supervisors' ratings were thus employed as criteria.

five recruiters mentioned most often as the best recruiters were used in the High Criterion Group, while the five recruiters mentioned most often as the poorest recruiters were placed in the Low Criterion Group (N = 45 for the High Criterion Group, and N = 43 for the Low Criterion Group).

The 4- to 6-hour battery was administered to the High and Low Criterion Groups in each DRC, and additional information concerning each recruiter's race, religion, and aptitude scores was collected from Army personnel files. Means for the keyed inventories and the aptitude scores were computed separately for the High and Low Criterion Groups. None of the personality measures or aptitude scores differentiated significantly between the two groups. Twenty background items pertaining to work habits, styles of handling finances and debts, educational background, and reaction to challenging or stressful situations did appear to distinguish the two groups; however, the significance levels were in general greater than .05 for these items.

Brown et al. suggested that these recruiters were a relatively homogeneous group, because all were required to meet several minimum qualifications (e.g., GCT scores, age, and rank) and had an average length of service of fourteen years. Because of this, the recruiters may have formed similar attitudes and opinions, and thereby limited the variance in attitude, personal preference, and personality inventory scores. The low validities, therefore, may have been due to restriction in range.

Graf and Bower (1976). Building on the Abrahams et al. work, Graf and Bower (1976) investigated the usefulness of the Strong Vocational Interest Blank and the Navy RIS scales for selecting Marine Corps recruiters. The SVIB was administered to each of 98 recruiters working in the Los Angeles and Santa Ana areas; of these, 77 inventories were completed and returned. Criterion measures consisted of officer performance ratings using a 3-point scale for below-average, average, and above-average recruiter performance.

Graf and Bower (1976) recognized that the Navy Recruiter Interest Scale (RIS) developed on a sample of Navy recruiters might also predict Marine Corps recruiter performance, and scored the Marine Corps recruiter responses on the RIS in order to evaluate the scale's validity. Results were in the predicted direction (a t-test significant beyond the .05 level) for a comparison of the upper and lower 35 percent of the RIS score distribution. In addition, the higher the RIS score, the greater the likelihood that the recruiter was in the effective group. Recruiters were then dichotomized as above-average or below-average according to the supervisory field ratings, and the two groups' RIS scores were compared. A significant difference (beyond the .05 level) in the predicted direction ( $r/pt. bi$ -serial = .30,  $p < .05$ ) was found.

Although the RIS effectively discriminated between above-average and below-average Marine Corps recruiters, Graf and Bower hypothesized that differences between Navy and Marine Corps recruiters might warrant a separate selection device. The authors elected to develop a Marine Corps Recruiter Interest Scale (MCRIS) key based on Marine Corps recruiter responses and administered the SVIB to a second sample of Marine Corps recruiters (N = 91) attending the Marine Corps recruiting school. Two years after the SVIB administration performance ratings for these recruiters were collected from supervisors. This criterion measure of

recruiter performance was used to dichotomize the combined sample ( $N = 168$ , consisting of 77 recruiters from the original sample and 91 Marine Corps students at the recruiter school) into below-average and above-average recruiters. SVIB items were then weighted according to the differential endorsement patterns of the two groups, and included in the MCRIS if endorsement differences between above- and below-average recruiters exceeded 15 percent. This procedure resulted in 75 items for the key. (Interestingly, only 13 of the 75 items in MCRIS are also found in the RIS.)

Responses from both of Graf and Bower's Marine Corps recruiter samples were scored together using the RIS and the MCRIS, resulting in validity coefficients of .67 for the MCRIS and .29 for the RIS. These two coefficients cannot be compared directly, however, because the RIS validity coefficient reflects the key's cross-validity, while the cross-validity of the MCRIS cannot be gauged from these data. Graf and Bower recommended assessing the MCRIS's utility as a selection device by evaluating the cross-validity of the key on a larger, more representative sample of Marine Corps recruiters. They also indicated concern for their criterion measure of recruiter performance and recommended attempting to develop a more reliable method of measuring recruiter performance.

The concerns of Graf and Bower (1976) mentioned above are similar to those of Abrahams et al. (1973) and point to some important criterion-related issues. Ideally, recruiter performance criteria should be uncontaminated by extraneous factors, should show consistent agreement by multiple raters on the performance of individuals, and should cover all aspects of recruiter performance. Although commanding officer nominations have proven useful, such single overall judgments of effectiveness/ineffectiveness cannot reflect individual differences in different aspects of performance. Thus, multiple facets of the recruiter job cannot be assessed. Additionally, officers may rely upon the reputations of their subordinate recruiters in making nominations or ratings instead of actual observation of recruiter performance.

Borman, Hough, and Dunnette (1976). Explicitly defining the behaviors leading to effective or ineffective performance is one approach that may focus raters on actual performance. Borman, Hough, and Dunnette (1976) employed such a strategy in developing behavior based rating scales for Navy recruiters. These authors used the behavior analysis methodology (Flanagan, 1954; Smith & Kendall, 1963) to develop eight 9-point behaviorally anchored rating scales: Locating and Contacting Qualified Prospects; Gaining and Maintaining Rapport; Obtaining Information from Prospects and Making Good Navy-Person Fits; Salesmanship Skills; Establishing and Maintaining Good Relationships in the Community; Providing Knowledgeable and Accurate Information about the Navy; Supporting Other Recruiters and the Command; and Administrative Skills. These scales, along with nine unanchored scales (including an Overall Performance Scale), were used to gather performance ratings employed as criteria in validation studies conducted by Borman and his colleagues (Borman, Rosse, & Abrahams, 1980; Borman, Rosse, & Rose, 1983; Borman, Toquam, and Rosse, 1979).

Borman, Toquam, and Rosse (1979). In this study, the authors developed an empirically keyed non-cognitive selection battery called the Special Assignment Battery (SAB) for enlisted Navy and Marine Corps recruiters.

Their research entailed five major steps: (1) criterion measures (rating scales and production measures) were developed; (2) ratings data were collected and analyzed; (3) a predictor battery was developed on the basis of hypothesized predictor-criterion relationships and was administered to recruiters; (4) data collected on the preliminary battery were used to develop new item pools for a second validation step; and (5) the final battery consisting of personality, vocational interest, and background items was employed in a concurrent validation study.

After developing Navy recruiter rating scales, Borman and his colleagues developed a revised rating version for Marine Corps recruiters. They collected peer, self, and supervisor ratings for 267 Navy recruiters and 118 Marine Corps recruiters and pooled the self-peer-supervisor ratings. The median interrater reliability was .57 for the pooled Navy recruiter ratings and was .48 for the pooled Marine Corps sample. A factor analysis of the ratings was then performed for the combined Navy/Marine Corps sample yielding a three factor solution: Selling Skills, Organizing Skills, and Human Relations Skills. A fourth criterion, Overall Performance, was added to serve as a summary criterion.

Borman et al. also gathered production data (the next section will discuss problems and limitations of production as a criterion) for each recruiter. This was defined as the number of accessions over the six-month period from May to October 1976. Recruiters' production data were standardized within recruiting district to provide a production index adjusted for geographic location. This production index, the scores on the three factors from the ratings data, and Overall Performance provided five criteria.

Intercorrelations among the five criteria yielded similar results for the Navy and Marine Corps samples. Overall Performance correlated with the Selling Skills, Human Relations Skills, and Organizing Skills criteria .79, .50, and .51 for the Navy and .72, .42, and .40 for the Marine Corps sample. The production index correlated .52 (Navy) and .59 (Marine Corps) with Overall Performance and .43 (Navy) and .45 (Marine Corps) with Selling Skills. The remaining correlations were .30 or below.

Borman et al. (1979) had selected non-cognitive items they hypothesized would predict recruiter performance in each of four performance categories--Overall Performance, Selling Skills, Organizing Skills, and Human Relations Skills. The selection of items was from a number of published instruments on the basis of data from an earlier pretest (N = 62 Navy recruiters). Pretest intercorrelations between the predictor items and the five criteria had been examined to identify pools of valid personality, interest, and biographical items for each of the four performance categories. The post-pretest battery, entitled the Special Assignment Battery (SAB), contained 310 personality items, the entire 325-item Strong-Campbell Interest Inventory, and a Biographical Survey. The authors also developed 12 keys for the SAB with four pools of personality items (one for each performance category), four pools of interest items, and four pools of biographical items.

The SAB was then administered to the 267 Navy and 118 Marine Corps recruiters for whom rating data were concurrently obtained. Correlations between the predictor keys and the five criteria were examined and cross-validated via a Monte Carlo procedure (separately for the Navy and Marine Corps samples).<sup>3</sup>

For both the Navy and Marine Corps samples, personality and vocational interest keyed scores were significantly correlated with the four performance criteria they were targeted to predict. These scores were not significantly related, however, to the production index. The biographical survey keys produced low validities for both types of criteria.

Data from the Navy sample ( $N = 267$ ) alone were then used to develop new composite SAB keys that were again cross-validated via the Monte Carlo procedure. This resulted in four composite SAB keys that combined personality, interest, and biographical items for predicting each of the performance criteria. The validities of these final composites against the four performance criteria are shown in Table 2.

Borman, Rosse, and Abrahams (1980). In finalizing and revising the SAB, these authors used the data described above to develop new items and cross-validate old items for the SAB and to investigate the personal characteristics tapped by each key. Their first step was to factor analyze (Navy sample,  $N = 267$ ) the items within the eight item pools (i.e., four sets of personality items and four sets of interest items) to identify underlying personality and interest constructs. The constructs identified in this analysis appear in Table 3. Next, they wrote new items targeted toward tapping these constructs. These new items, along with items found previously (Borman et al., 1979) to be valid, formed a new SAB which included four keys for personality items and four keys for interest items (i.e., one for each performance construct).

---

<sup>3</sup>Most methods for estimating the cross-validity of predictors involve four steps: (1) dividing the total sample into two subsamples at random (a developmental sample and a cross-validation sample); (2) developing a prediction equation using the developmental sample; (3) using the interim prediction function to compute point predictions for each subject in the cross-validation sample; and (4) computing the appropriate measure of association that is considered the estimate of validity (usually a Pearsonian coefficient). Because the validity estimated in this manner is conditional upon a particular "split" or randomization of the total sample, Borman et al. (1979) employed a Monte Carlo technique that repeats steps 1-4 above, randomizing samples and reestimating the prediction function with each repetition. The standard error of the mean within-sample distribution was computed with each repetition. When the standard error was sufficiently reduced, repetitions were stopped. This process gave rise to a distribution of "cross-validities," the mean of which was taken as the estimate of validity.

Table 2  
SAB Composites/Keys (Pooled) Validity Estimates

	Performance Criteria			
	Selling Skills	Human Relations Skills	Organizing Skills	Overall Performance
Navy <sup>1</sup>	.24**	.17**	.31**	.22**
Marine Corps <sup>2</sup>	.22*	.22*	.38**	.27**

<sup>1</sup>N = 267

<sup>2</sup>N = 118

\*  $p < .05$

\*\*  $p < .01$

The new SAB was then administered to a new sample of 194 Navy enlisted recruiters. Again, peer and supervisory ratings and production data were gathered as criteria. Ratings were obtained for 17 rating dimensions including eight Behavior Summary Scales, Overall Performance, and eight other scales. Interrater reliabilities for the pooled peer and supervisor ratings ranged from .34 to .77 with a median of .51. A factor analysis of the pooled ratings yielded a three factor solution--Selling Skills, Human Relations Skills, and Organizing Skills--which was very similar to the solutions generated previously (Borman et al., 1979).

Because new items targeted toward tapping specific constructs were included on the SAB, Borman et al. first computed intercorrelations between old and new predictor construct composites of items to ensure that the items were indeed measuring the constructs they were intended to measure. For personality items, the median correlation was .56 ( $p < .001$ ), and for interest items the median correlation was .67 ( $p < .001$ ), indicating that the new items did, in fact, measure the target constructs.

The authors then computed validities of the old item composites, the new item composites, and composites of old-plus-new items against the supervisor/peer ratings criteria. For each performance category (e.g., Selling Skills) a single SAB composite was formed across the personality items targeted toward that criterion, and another composite was formed for the interest items. The old-plus-new personality item composites correlated .23 ( $p < .01$ ) with Selling Skills, .24 ( $p < .01$ ) with Human Relations Skills, .15 ( $p < .05$ ) with Organizing Skills, and .22 ( $p < .01$ ) with Overall Performance. Similarly, the old-plus-new interest item composites correlated significantly ( $p < .01$ ) with Selling Skills ( $r = .21$ ), Human Relations Skills ( $r = .22$ ), and Overall Performance ( $r = .22$ ).

Table 3  
Factors Identified for Each Performance Category

Performance Categories	Factor/Construct
	Personality Items
Selling Skills	<ol style="list-style-type: none"> <li>1. Good impression.</li> <li>2. Impulsive, carefree <u>vs.</u> order, planning ahead, systematic, level-headed.</li> <li>3. Enjoying being center of attention, leading, showing off, and speaking before a group.</li> <li>4. Working hard and with confidence, being happy <u>vs.</u> being unhappy, giving up easily, disgruntled about life.</li> </ol>
Human Relations Skills	<ol style="list-style-type: none"> <li>1. Preference for working with and being with people.</li> <li>2. Spontaneity, impulsivity, "fast and careless," rebellious, tendency to have bad moods.</li> <li>3. Unhappy, lack of confidence, disgruntled about life.<sup>a</sup></li> <li>4. Ambitious, working hard, pushing self.</li> </ol>
Organizing Skills	<ol style="list-style-type: none"> <li>1. Order, planning ahead, well organized <u>vs</u> impulsive, acting without thinking, "fast and careless."</li> <li>2. Leading and influencing others, giving orders, demanding of self, ambitious, dominant.</li> <li>3. Unhappy, discouraged, doing little in life, giving up hope, feeling useless.</li> <li>4. "Bad actor," was unruly and rebellious in school, unsocialized.</li> </ol>
Overall Performance	<ol style="list-style-type: none"> <li>1. Doing more than expected <u>vs.</u> giving up, working just hard enough.</li> <li>2. Impulsive, "fast and careless" <u>vs.</u> order, methodological, planning ahead.</li> <li>3. Leading and influencing others, dominant, strong personality.</li> <li>4. Good impression <u>vs.</u> admitting occasional meanness, grouchiness, disgust with self, discouragement, uselessness, bad mood.</li> <li>5. People oriented, liking to be around others and close to others, open to other people.</li> </ol>
Vocational Interest Items	
Selling Skills	<ol style="list-style-type: none"> <li>1. Interest in extroverted, dominant, leadership activities and occupations.</li> <li>2. Interest in occupations involving attention to detail.<sup>a</sup></li> <li>3. Interest in law and politics.</li> <li>4. Interest in sports and competitive activities.</li> </ol>
Human Relations Skills	<ol style="list-style-type: none"> <li>1. Interest in dominant, extroverted, social activities.</li> <li>2. Interest in teaching and counseling.</li> <li>3. Interest in "feminine" occupations and activities.</li> <li>4. Interest in newspaper reporting and foreign service.</li> <li>5. Interest in sports and competitive activities.</li> <li>6. Interest in religion and in being around the sickly.</li> </ol>
Organizing Skills	<ol style="list-style-type: none"> <li>1. Interest in politics and high level management jobs.</li> <li>2. Interest in bookkeeping, statistical, and detail work.</li> <li>3. Interest in "feminine" occupations and activities.<sup>a</sup></li> <li>4. Interest in leadership and responsibility.</li> </ol>
Overall Performance	<ol style="list-style-type: none"> <li>1. Interest in law and politics, and management occupations and activities.</li> <li>2. Interest in activities and occupations that require extroversion, dominance, responsibility, and leadership.</li> <li>3. Interest in sports and competitive activities.</li> <li>4. Interest in teaching and counseling.</li> <li>5. Interest in "feminine" occupations.<sup>a</sup></li> </ol>

<sup>a</sup>These constructs related negatively to their target performance criteria.

To develop a final SAB, the authors examined the cross validities for the old SAB personality and interest items and selected items that demonstrated consistent validity in both samples ( $N = 267$  and  $N = 194$ ). New items demonstrating good validity in the second sample ( $N = 194$ ) were also included in the final SAB. One hundred and ten personality and sixty interest items formed four keys (e.g., one key for Selling Skills which included personality and interest items, one key for Human Relations Skills which included personality and interest items, etc.). Correlations between the final SAB scores and pooled supervisory and peer ratings for the second sample ( $N = 194$ ) were .43 with Selling Skills, .46 with Human Relations Skills, .40 with Organizing Skills, and .43 with Overall Performance. The SAB scores were also correlated with each recruiter's raw production data. These validities were .22, .23, .13, and .26 respectively; three of which were significant ( $p < .01$ ).

In summary, results from research by Borman and his colleagues have shown that non-cognitive variables are useful predictors of both recruiter performance (supervisor ratings) and effectiveness (production). In addition, two aspects of this work warrant special attention. First, ratings on the 17 scales used in this research were factor-analyzed using three separate samples, and all analyses yielded similar three-factor solutions. Thus, it appears that the ratings obtained in these studies were tapping three relatively stable constructs underlying Navy recruiter performance. Second, although several other authors have developed empirically keyed predictors (e.g., Abrahams et al., 1973), they have not investigated the individual differences constructs being tapped by such predictors. In contrast, the Borman work represents a mix of both empirical and construct validation. (Because the SAB has been employed in several studies, we have briefly summarized relevant research in Appendix A.)

Borman, Rosse, and Rose (1983). In an effort to predict Navy officer recruiter performance, the SAB was again employed.<sup>4</sup> Borman, Rosse, and Rose (1983) examined the SAB items, evaluated the appropriateness of the item content for predicting officer recruiter performance, and wrote new items targeted toward constructs identified in the earlier research. Items from the SAB keys developed for enlisted recruiters (i.e., Selling Skills, Human Relations Skills, Organizing Skills, and Overall Performance) were used to form four corresponding officer recruiter composites. These composites were then correlated with supervisor ratings to assess the validity of these measures for predicting Navy officer recruiter success.

The predictor battery was administered to 132 commissioned officer recruiters with three months or more recruiting experience. Self and supervisory ratings were also gathered for each of the 132 recruiters on 9-point behaviorally anchored rating scales and an Overall Performance scale. The median intraclass correlation between supervisor and self-ratings on the behavior summary scales was .36, and on the overall performance dimension the intraclass correlation was .49.

---

<sup>4</sup>Navy officer recruiters are commissioned officers whose job involves recruiting officers.

Validity estimates were computed separately for the personality and vocational interest scales. Three of the four personality keys--Human Relations Skills, Administrative Skills and Overall Performance--correlated significantly with the performance criteria (.30, .18, and .19 respectively). Also, two of the interest keys--Selling Skills and Overall Performance--correlated significantly with the criteria (both  $r$ 's = .16).

The authors interpret the results of their work with Navy officer recruiters as indicating that:

"successful officer recruiters tend to be hardworking and ambitious, and they push themselves hard. They are also strongly oriented toward people activities. They like to be close to others, and they can be spontaneous and fun-loving in these social activities. Related to vocational interests, effective officer recruiters are especially interested in jobs and activities that require extraversion, dominance, responsibility, and leadership (p. 10)."

Reviewing the results of their work across studies, Borman and his colleagues note that there are similarities and differences between patterns of validities for officer and enlisted recruiters. In the personality realm, the hard work, and high ambition theme is present in both groups of successful recruiters. Additionally, successful enlisted and officer recruiters are characterized by spontaneity, even some of the negative aspects of it such as liking a "fast and careless" lifestyle and having a tendency to be a bit rebellious. Effective officer recruiters are more likely to have a strong people orientation, including a liking to be with and be close to others. Conversely, successful enlisted recruiters are characterized by an enjoyment of being the center of attention, even of "showing off". The more successful officer recruiters apparently project a more subtle people orientation.

Regarding vocational interests, a strong orientation toward social activities and toward occupations that require an outgoing and dominant self-presentation is characteristic of both officer and enlisted recruiters. Interests that describe successful enlisted recruiting personnel, but not so much effective officer recruiters, relate to sports and competitive activities and toward the law and politics. Successful officers have greater orientation toward leadership and responsibility than do successful enlisted recruiters.

In summary, a number of recruiter selection studies employed supervisor ratings or commanding officer nominations as criteria for assessing the validity of predictor measures. Non-cognitive predictors have proven successful in several studies. The work of Abrahams et al. (1973), Borman et al. (1979), Graf and Bower (1976), and Wollack and Kipnis (1960) suggests that vocational interests may successfully predict military recruiter performance. Studies employing personality measures as predictors (cf. Borman et al., 1979; Krug, 1972) found reasonably high relationships between these predictors and recruiter performance. Background or biographical measures have also predicted recruiter performance with some success (Borman et al., 1979; Krug, 1972). Considering all the work we have reviewed, a significant relationship between non-cognitive predictors and recruiter performance appears to be a consistent finding.

Cognitive predictors, on the other hand, have met with little success in predicting recruiter performance. Intelligence and aptitude measures have shown virtually no relationship with recruiter performance. Only verbal measures have yielded even moderate relationships with performance criteria (cf. Brown et al., 1975; Wollack & Kipnis, 1969). Overall, the cognitive predictor-recruiter performance (i.e., measured by ratings or nominations) relationship appears to be low and of little utility for recruiter selection research.

## Productivity as a Measure of Recruiter Effectiveness

Because recruiter production (i.e., gross number of recruits over a defined time period) is essentially the "bottom-line" criterion in recruiting and is the primary factor used by the armed services to evaluate recruiter success, identifying variables that might predict recruiter productivity is of particular importance. Unfortunately, the usefulness of production criteria in selection research is limited by two problems. First, objective production data are sensitive to environmental factors beyond the recruiter's control. Second, production data probably fail to tap all facets of the recruiter job. For instance, production data may not reflect recruiters' efforts to maintain rapport with the community or to provide support to other recruiters.

Research employing production criteria has proceeded in two areas: (1) research using production levels as criteria for validating predictors; and (2) research investigating environmental contaminants of production. We have, therefore, divided this chapter into two sections. In the first section we summarize research employing production as a primary criterion for predictor validation, and we discuss research focusing on factors influencing production in the latter section.

### Production as Criteria for Validating Predictors

Brogden and Taylor (1949). In an attempt to predict Army recruiter effectiveness, Brogden and Taylor (1949) developed a battery consisting of four paper-and-pencil predictors. These predictors included a measure of vocational interests and three measures tapping interests, hobbies, and background. The four predictor scales were then item-analyzed against a production criterion consisting of the average number of recruits brought into the Army by each recruiter per hour on recruiting duty. The resulting empirically keyed predictors were cross-validated using a sample of 475 recruiters, yielding a validity coefficient of .18.

Brogden and Taylor postulated that this low validity was due to the unreliability of the production criterion and to the contamination potentially present in such global objective indices. The authors reanalyzed the data using a turnover criterion. When this criterion was used, the summed scores of the four inventories yielded a biserial validity coefficient of .36.

The focus of the Brogden and Taylor study was on the criterion employed to validate predictors for use in selecting Army recruiters. The initial production criterion was criticized because it was thought to contain many factors not associated with recruiter effectiveness. Of course, the turnover criterion may also reflect factors unrelated to recruiter effectiveness. The turnover criterion did, however, result in a higher validity when it was correlated with the predictors. This increase in validity may be due in part to the relatively low reliability and significant contamination associated with the production criterion. It is also possible that measures such as those used here are simply better predictors of turnover than they are of recruiter production.

Larriva (1975). In an effort to overcome problems related to the gross production criterion, Larriva developed an adjusted production index for use in a concurrent validity study. All recruiters from a single Marine Corps District ( $N = 470$ ) were administered the predictor battery which included:

1. The 16PF--a personality test designed to measure an individual's personality in terms of 16 basic factors.
2. A 25-item experimental supplement designed to measure motivational distortion (a lie scale) and the strength of motivation to succeed as a recruiter.
3. Seven biographical items yielding information about years of service, age, sex, marital status, number of dependents, years of formal education, and population of subject's home of record.

The Recruiting Command made arrangements to record the number of Non-Prior Service (NPS) accessions for each recruiter throughout the year 1974. This was the initial production criterion used in the research.

After all the completed inventories were returned and the first quarter accession data were reported, Larriva analyzed the data using a regression formula developed by the Navy to predict recruiter performance (cf. Krug, 1972). This analysis indicated that the Navy formula was not a valid equation for predicting gross Marine Corps recruiter production. Larriva conjectured that the accessions criterion might not provide relevant effectiveness information; thus, he generated several effectiveness indices on the basis of assumptions about and corrections for the recruiting environment and geographical differences in production.

Using first and second quarter accession data, each production index was employed as the criterion measure and correlated with the personality, demographic, and experimental variables of the predictor inventory. Larriva then examined the resulting predictor-criterion relationships and selected the criterion index that yielded the most valid multiple correlation coefficient. The index selected, referred to as P16, indicated the number of accessions for urban and rural recruiters separately and corrected for geographical differences in relative effectiveness of recruiters. Accordingly, the validation sample was split into an Urban Recruiter Group ( $N = 308$ ) and a Rural Recruiter Group ( $N = 162$ ), and a weight correcting for geographical differences was applied to production scores. Larriva also developed subsets of these two samples of urban and rural recruiters that contained only the best and poorest recruiters. These two reduced samples were referred to as the Urban Hi-Low ( $N = 122$ ) and the Rural Hi-Low ( $N = 96$ ) groups. The P16 criterion was then used as the effectiveness measure (criterion) and stepwise multiple regression was utilized to generate four formulae for predicting recruiter success from the inventory data (i.e., one for each of the two full samples--Urban and Rural, and one for each of the subset samples--Urban Hi-Low and Rural Hi-Low).

The means and standard deviations of the predicted and actual scores were calculated for the four data bases and were used to establish cutoff

scores. These cutoff scores were then used to assess the effectiveness of the prediction equations and to select or delete recruiters in a cross-validation sample. The cross-validation sample consisted of 98 recruiters who were not included in the validation sample but who had a minimum of three months accession data. As prescribed by the production index, the sample was split into an Urban Recruiter Group ( $N = 66$ ) and a Rural Recruiter Group ( $N = 32$ ), effectiveness scores (the P16) were generated for each recruiter, and subsets of the two full samples (i.e., Urban and Rural) were again developed.

The two urban prediction equations were applied to the urban cross-validation sample and results of the prediction equations were compared according to accuracy and expected increase in productivity. The same procedures were followed to compare the results of the Rural prediction equation and Rural Hi-Low prediction equation when applied to the rural cross-validation sample. Results for the two comparisons are listed in Table 4.

Table 4  
Comparison of Rural and Urban Recruiter Samples  
After Application of Regression Equations

	<u>Urban</u>	<u>Urban Hi-Low</u>	<u>Rural</u>	<u>Rural Hi-Low</u>
Number deleted by cutting score	24	34	12	12
Percent correct cutoff	88%	76%	75%	83%
Mean production score of those deleted	9.29	9.91	10.06	9.02
Mean production score of those remaining	12.29	12.54	12.39	13.02
Expected increase in productivity	10%	12%	7.6%	13%

From these comparisons, Larriva concluded that the Urban predictor equation and Rural Hi-Low equation would prove successful in selecting Marine Corps recruiters. The author also suggested that both equations could be used to analyze each predictor inventory, thus allowing the Marine Corps Recruiting Command to select the optimal setting, urban or rural, for each new recruiter.

In summary, when recruiting setting (urban-rural) and geographical differences were accounted for in the criterion, relatively successful prediction equations were formed for urban recruiters and for rural re-

cruiters. Cross-validity results suggest that the predictor inventory may be a useful selection device for screening candidates for the Marine Corps recruiter job. The Navy prediction equation did not effectively predict unadjusted Marine Corps recruiter production. The original Navy equation was developed, however, against commanding officer nomination criteria rather than against production. Thus, the low cross-validity may indicate that the Navy equation is more useful in predicting performance criteria than production. It would be of interest to apply the Navy prediction equation for the Larriva inventory to the cross-validation sample using the P16 criterion. Apparently, Larriva never correlated the Navy prediction equation scores with scores on the P16 criterion.

A final point of interest in this research has to do with the selection of a criterion. Larriva essentially selected the criterion that correlated most highly with the predictors. A clearly more acceptable (and justifiable) method is to define the criterion, or criteria, as precisely as possible and then to select a measure that provides relevant and reliable measurement of each criterion without regard to the predictors. The cross-validation procedure employed in Larriva's study makes this criticism less severe, but the P16 criterion would have been better justified as a measure of effectiveness if predictor data had not dictated its choice.

Atwater, Abrahams, and Trent (1986). Recently, Atwater et al. (1986) have conducted two studies employing production data as criteria for validating the Special Assignment Battery (SAB) (cf. Borman et al., 1983): (1) a large-scale concurrent validation study; and (2) a predictive validation study.

For the concurrent study, these authors administered the SAB to 1,005 Marine Corps recruiters who had been on the job for at least six months. Two criterion measures were obtained for each recruiter - gross productivity (average number of monthly contracts written) and a supervisor rating on a three-point overall performance scale.

Recall that the SAB is a non-cognitive predictor battery which includes four scales - Selling Skills, Human Relations Skills, Organizing Skills, and Overall Performance. Correlations between the predictor scales and gross productivity were .23 ( $p<.01$ ), .15 ( $p<.01$ ), .09 ( $p<.05$ ), and .24 ( $p<.01$ ), respectively. Correlations between the predictor scales and the overall performance rating were .19, .15, .13, and .20, and were all significant ( $p<.01$ ). The authors compared these results to those obtained by Borman et al. (1983) on a sample of Navy recruiters ( $N = 194$ ) and found that the patterns of correlations were quite similar and that there were no statistically significant differences between the production/predictor correlations obtained in the two studies.

For the predictive study, the SAB was administered to 664 Marines who were being screened for assignment to recruiter duty during 1981 and 1982 (SAB scores were not used in actual assignment to recruiting duty). Four hundred and thirty-one of the 664 Marines were assigned to recruiting duty by 1983 and were tracked through recruiting school and for at least a year of recruiting duty.

Two criteria were employed for validating the predictor battery - production and completion of tour. The production index was net monthly

production computed by subtracting the number of recruits who left the service while in the delayed entry program or during basic training from each recruiter's gross production. The completion of tour measure indicated whether the recruiter had completed at least one year of recruiting duty or had been removed from recruiting duty.

Of the 431 Marines assigned to recruiting duty, 76 failed to complete recruiting school, and criterion data were not available for 40, reducing the sample size to 315. Correlations between the four predictor scales and the production measure were .27 (Selling Skills), .17 (Human Relations Skills), .21 (Organizing Skills), and .29 (Overall Performance). Again, these correlations were consistent with those found previously in research with this predictor battery.

A summary of the attrition rates appears in Table 5. The authors computed a composite of the four SAB scales, the recruiter potential composite (Recpot), and compared Recpot scores for the samples at various stages in the study. As shown in Table 5, the average Recpot score for the 76 Marines who failed to complete recruiting school was lower than that of those completing school ( $N = 355$ ); the difference, however, was not statistically significant. Of the 315 recruiters who were located after recruiting school, 84 did not have normal rotation to their tour (i.e., 73 were relieved for various reasons, and 11 completed their active service in the Marine Corps before the year ended). The mean Recpot score for these 84 recruiters was nearly one-half standard deviation below the mean of those who completed their tour. Although the eleven recruiters who ended their active service were not dismissed for poor performance, their production was significantly below the average for Marines who completed their tour (1.8 recruits per month vs. 2.3, s.d. = 1.0).

In sum, the non-cognitive predictor scales (SAB) developed using a sample of Navy recruiters successfully predicted Marine Corps recruiter productivity in a concurrent study including over a thousand Marine Corps recruiters and in a predictive study ( $N = 315$ ). The authors, therefore, noted that the similarity of these results suggests "stability over time, as well as generalizability of SAB scores across services (p. 3.)."

#### Environmental Contaminants of Production

Because environmental factors as well as recruiters' personal characteristics (e.g., personality, interests) may influence productivity, a number of studies have been conducted to delineate factors outside the individual's (recruiter's) control that might contaminate production data. The paragraphs below summarize research in this arena.

Bennett and Haber (1973). To investigate factors influencing recruiter production, these authors collected data on several environmental and background variables from 259 Marine Corps recruiters. The recruiters represented 29 recruiting stations throughout the United States. The information was collected from performance records and in a survey of recruiters and their jobs. Bennett and Haber divided the variables into three categories:

Table 5  
Means, Standard Deviations, and Sample Sizes  
for the Predictive Recruiter Study<sup>1</sup>

Sample	<u>Recpot<sup>2</sup></u>	
	Mean	SD
Of 644 applicants:	58.5	45.2
431 were assigned to recruiting school	60.4	48.4
213 were not assigned	55.0	47.2
Of 431 assigned to recruiting school:		
355 completed school	61.7	48.3
76 failed to complete	54.4	49.9
Of 355 who completed school:		
315 were located in field	61.6	48.1
40 were not located	62.5	47.4
Of the 315 located in field:		
231 had normal tour completion	66.4	44.6
84 did not have normal tour completion	48.6	53.6

<sup>1</sup> Atwater, Abrahams, & Trent (1986).

<sup>2</sup> Recruiter potential composite (Recpot); score range: -75 to +193.

1. Selection variables--General Classification Test scores; age; race; level of education; number of dependents; previous service as a career planner (i.e., counselor) or drill instructor; method of assignment to recruiting duty (volunteer/assigned); and opinion about recruiting duty as a financial hardship.
2. Deployment variables--whether the recruiter is assigned in or out of his or her home state; distance from home state; type of area assigned to--urban/suburban or rural; number of times assigned; hours per week spent on recruiting; and percentage of time spent out of the office recruiting.
3. Evaluation variables--number of months the individual has been on his or her current tour of duty and the recruiter's percentile rank in Marine Corps Recruiter Class.

Gross productivity (i.e., the average number of recruits enlisted per month) was employed as the measure of effectiveness for each recruiter; yet Bennett and Haber noted that differences in gross productivity cannot be completely attributed to individual recruiter differences in performance. Differences may be due in part to a variety of regional differences such as the number of qualified prospects or local attitudes toward the military service. For this reason, the sample was divided into two groups on the basis of the average regional enlistment rates for 1971. The first group contained those recruiters from high enlistment recruiting stations ( $N = 109$ ) and the second group contained recruiters serving in low enlistment areas ( $N = 150$ ).

The authors used multiple regression to determine the magnitude of the relationship between each variable and recruiter productivity, and computed a separate regression equation for each enlistment area group. In the high enlistment area group, three variables were found to be significantly related to productivity. The first variable, type of area, showed that urban/suburban recruiters enlisted more persons per month ( $p < .01$ ) than rural recruiters. The second variable, geographical assignment, showed that recruiters stationed in their home state enlisted more persons per month ( $p < .01$ ) than those stationed more than 500 miles away from their home state. The third indicated that productivity was negatively related to increased reassessments ( $p < .05$ ). These three variables accounted for 35 percent of the production variance.

In the low enlistment areas, two variables correlated significantly with productivity. The first significant variable indicated that those recruiters who feel recruiting duty is a financial hardship enlist more persons per month ( $p < .05$ ) than those who do not. The second significant variable suggested that those recruiters who had previously served as career planners enlisted a greater number of persons per month ( $p < .05$ ) than did those who had not been career planners. These two variables account for only 12 percent of the productivity variance, however, and it should be noted that the regression equations obtained in the study were not cross-validated. Thus, the relationships found here may well overestimate the true usefulness of these variables.

Brown, Wood, and Harris (1975). These authors also sought to develop a valid criterion measure of recruiter effectiveness by focusing on an index

used unsuccessfully in previous recruiter studies--total production scores. Brown et al. noted that these scores were contaminated by "opportunity bias" arising from territorial factors that fall outside the control of the individual recruiter. They hypothesized that if territorial factors were somehow partialled out of the production score variance, the result would be a more realistic, unbiased picture of recruiter effectiveness.

Brown and colleagues moved in this direction first by identifying territorial variables that could potentially bias production figures. A total of twelve such biasing agents were noted, including the average number of enlistments per recruiter in the recruiter's District Recruiting Command or DRC; the proportion of territory that is metropolitan, suburban, or rural; and the recruiters' length of experience. A nationwide random sample of Army recruiters was then developed by requesting each commander from the five Regional Recruiting Commands (RCRs) to supply names of 100 recruiters from their respective RCRs. Commanders selected recruiters randomly according to their social security numbers, and also supplied total production figures for each recruiter for the period from July to December 1973. Territorial information from each DRC represented by recruiters in the sample was then compiled by Army personnel from available records and from a special market survey they conducted.

Scores on the twelve territorial factors were subjected to a stepwise multiple regression with total production scores serving as the criterion. Three variables contributed significantly ( $p < .05$ ) to the prediction of production scores: average production per recruiter in subject's DRC, proportion of all enlistees in the DRC who chose the Army (rather than another service), and suburban proportion of the zone were positively related to production. These three variables accounted for 51 percent of the production score variance.

Next, the three territorial factors were used to develop a multiple regression equation that yielded predicted production scores for each recruiter. These predicted scores were then employed to compute unbiased production scores obtained by computing the ratio of total production to predicted production scores and multiplying by 100. The authors termed the resulting scores as Benchmark Achievement Scores (BAS), which were corrected for DRC production, proportion of young persons joining the Army in the DRC, and the proportion of the DRC that is suburban. These adjusted scores should provide relatively unbiased production information.

Brown et al. also developed adjusted scores that are easier to compute than BAS; these scores were termed Simple Achievement Scores (SAS). They were derived by computing a ratio of a recruiter's total production to the average production in the DRC. Because this ratio was the variable that accounted for most of the variance in the multiple regression equation, Brown et al. recommended that the Simple Achievement Scores be used as future criterion measures. Simple Achievement Scores also correlate highly with the BAS ( $r = .96$ ). The development of these adjusted production measures was, however, not complete when these authors conducted a validation study and they were, therefore, not employed as criteria.

Arima (1977). Similar to the Brown et al. (1975) study, Arima (1977) employed a regression approach to identify environmental variables associated with Navy recruiter productivity. Arima collected data on several

educational environment variables such as average pay for local high school teachers and average daily attendance in local high schools. He then scored 268 recruiters in three recruiting districts on these variables, as well as two dummy variables representing district membership. The environmental variables were entered into a step wise multiple regression analysis and results showed that approximately one-third of the variance in recruiters' raw production could be accounted for by the variables studied.

A potential problem with both the Arima (1977) and Brown et al. (1975) studies has to do with the distinction between district membership as a variable and more specific environmental variables such as employment rate and median family income. District membership has consistently been shown to be a key factor in recruiter production, yet this finding says nothing about which aspects of different districts (or which environmental variables) result in the recruiter production effects. That is, environmental factors provide possible reasons for between-district differences in production levels. Thus, for conceptual reasons, it appears appropriate to keep separate consideration of specific environmental variables and district membership. The Arima and Brown et al. analyses fail to make this distinction, in fact including both district membership and environmental variables in the same regression equations.

Borman, Rosse, and Toquam (1982). To delineate the environmental variables impacting on recruiter production, these authors developed a list of potentially important factors from interviews with officers and Navy Recruiting Command officials. They obtained seven factors including:

1. Unemployment rate in the recruiting area.
2. Level of competition (i.e., the ratio of number of other service recruiters/number of Navy recruiters).
3. Ratio of military to civilian pay in an area.
4. Ratio of military to total population in an area.
5. Ratio of Department of Defense civilian employees to total population in an area.
6. Propensity of young people to enlist in an area (i.e., data from the Youth Attitude Tracking Survey).
7. Ratio of support personnel available to process recruit data/the total number of recruiters.

Measures of these variables were obtained at the district level ( $N = 43$ ), and mean production scores were computed for these districts.

To examine relationships between the environmental variables and production, the authors correlated each variable with mean district production. Of the seven environmental factors, three correlated consistently with mean district production across four quarters of production data (Military/Civilian Pay Ratio,  $r = .36$ ; Unemployment Rate,  $r = .30$ ; and Propensity to Enlist,  $r = .23$  against FY 80 total production), indicating that district level production is affected by some of the environmental factors.

To assess the impact of the environmental variables (in combination) on individual recruiter production levels, the authors employed an ANOVA procedure producing estimates of variance accounted for by the environmental variables, by differences between districts, and by within-district factors. Although the effects of the environmental variables on individual recruiter production levels was significant ( $F = 10.87$ ,  $p < .001$ ), they accounted for only five percent of the total variance in production. The total variance accounted for by differences between districts was 7.4 percent (after the effects of the environmental measures were removed).<sup>5</sup> Importantly, most of the variance in recruiter production (87.6%) was due to within-district factors. In other words, the bulk of the variance in recruiter production may result from between-recruiter differences in recruiting ability or, perhaps, environmental factors at more fine-grained geographical levels (e.g., zone, station).

In sum, a number of studies have investigated either the personal characteristics that might predict production or the environmental factors that may influence individual recruiter productivity. With regard to person characteristics, non-cognitive variables have validly predicted production criteria (cf. Atwater et al., 1986; Bennett & Haber, 1973; Borman et al., 1980; Larriva, 1975). Although they have also shown lower validities than when ratings are used as criteria (cf. Borman et al., 1979; Brogden & Taylor, 1949), recent large-scale validation research evaluating non-cognitive predictors against production supports this relationship.

Because environmental factors, as well as the skills and abilities of individual recruiters, influence recruiter productivity, research has focused on identifying contaminants of production data. Brown et al. (1975) and Larriva (1975) developed production indices adjusted to account for territorial factors. Researchers have also worked toward defining the environmental variables impacting production (Arima, 1977; Borman et al., 1982; Brown et al., 1975). Variables describing the recruiting setting (e.g., unemployment rate, military/civilian pay ratio) have produced significant correlations with production data. To date, research in this arena has focused on variables influencing production at the district level (approximately equivalent to battalion level). Adjusting production indices for district, or battalion, differences does remove some contaminating variance; however, there appears to be considerable production variance within districts (Borman et al., 1982). Further research at a more fine-grained level (e.g., the company level) is needed before purer production indices can be developed.

---

<sup>5</sup>Forty-one percent of the total variance between-districts in production was accounted for by the seven environmental factors.

## Predicting Recruiter Success in Training

Studies discussed in the preceding chapters employed subjective performance criteria (i.e., ratings) or an objective effectiveness criterion (i.e., production) to define on-the-job success as a recruiter. In this section, the studies discussed focus on criteria for recruiter success in training. These criteria can be subjective or objective and performance or effectiveness measures. Yet training criteria are different conceptually, and often methodologically, from measures of success on-the-job. Conceptually speaking, training performance may include behaviors such as studying or practicing that are not highly related to on-the-job performance. Likewise, behaviors important for recruiter performance on-the-job (e.g., actions taken to establish rapport with the community) may not be highly related to those important for successful performance during training. Thus, with respect to training criteria, the term "performance" may imply different performance requirements than is the case with job performance.

Massey and Mullins (1966). In this early study, a dichotomous, objective criterion, graduation versus elimination from training, was employed to assess the validity of an Air Force recruiter selection battery. The authors constructed an eight-inventory battery designed to measure qualities such as empathy, surgency (friendliness and sociability), and perseverance, all thought to be desirable in recruiters. The measures are listed below:

1. Airman Preference Test--a 43-item five-choice test designed to measure empathy.
2. CEI Opinion Questionnaire--83 items describing a person's behavior, an individual's opinions, old proverbs, and expressions of interest. The examinee is to show agreement, disagreement, or no opinion. Three scores (Complexity, Ego Strength, and Introversion) were considered likely predictors of recruiter success.
3. Community Information--43 items to gain factual information about areas of former residence. This was designed on the rationale that a person who knows more about the community is one who is more interested in it and consequently one who should be a more effective recruiter in that location.
4. Descriptive Adjective Inventory--182 paired adjectives used to describe a person's general appearance, typical behavior, and relations with others. The examinee selected the one adjective of the pair that best describes him or her. Respondents were scored on three scales (Surgency, Cooperativeness, and Orderliness).
5. FCSRI-A--190 items, arranged in pairs of statements, that refer to a person's appearance, typical behavior, and relations to others. Three scores (Surgency, Cooperativeness, and Orderliness, the same three variables as measured by the Descriptive Adjective Inventory) were considered likely predictors of recruiter effectiveness.

6. Recruiter Language--made up of four subtests:

- a. 15 five-choice items--examinee must select the initial letter of a missing word. The definition is provided.
- b. 15 five-choice items--examinees must select the correct arrangement of words appearing in progressive order. The words express varying degrees of the same general concept.
- c. 10 five-choice items--examinee must select the best sentence with respect to grammar and usage.
- d. 15 five-choice items--a sentence is provided; an examinee must select the interpretation that means most nearly the same as the sentence.

7. Texas Social Insight--40-item test to describe ordinary, unusual, or embarrassing situations. The examinee is to select the alternative he/she sees as the correct course of action.

8. Word Power--test of ability to call to mind certain words. At a given signal the examinee records as many words as possible beginning and ending with a given letter of the alphabet.

In a preliminary validation step, the battery was administered to 210 recruiter students prior to entry in the recruiter course. Twenty-four scores, derived from the eight inventories, were correlated with the dichotomous criterion measure of graduation or elimination from the recruiter course (178 graduated and 32 were eliminated). Four of the scores--Recruiter Language, Community Information, FCSRI-A (Surgency), and the Texas Social Insight Test--yielded significant point biserial correlations ( $p < .01$ ).

A combination of the Recruiter Language, Community Information, and FCSRI-A (Surgency) scores yielded a multiple correlation of .23 ( $p < .01$ ). (The Texas Social Insight Test did not add significantly to the multiple correlation.) The three inventories, with scoring weights of 2, 1, and 1 respectively, were employed as an interim selection battery, pending further validation of the eight-inventory composite.

Next, the full eight-inventory battery was administered to a larger sample of recruiter students ( $N = 1067$ ) prior to entry in the recruiter course. Background information including age, education, marital status, and number of dependents was also collected. As the criterion, pass/fail information was subsequently obtained for each recruiter student.

The validation sample ( $N = 965$ , students with complete information) was randomly split into two groups--Sample B ( $N = 485$ ), the validation sample; and Sample C ( $N = 480$ ), the cross-validation group. Twenty-four scores, derived from the eight inventories, were generated for Sample B recruiters. These scores, along with the background variables, were correlated with the recruiter course pass/fail criterion. No combination of predictors significantly ( $p < .05$ ) increased prediction over that obtained using the three scores from the interim battery. The scoring weights of the interim battery (RSST-63) were altered a number of times and used to score Sample B

recruiters. These scores were validated against the pass/fail criterion and again, no combination of regression weights surpassed the prediction capabilities of the interim battery with the original weights ( $r = .34$ ,  $p < .01$ ). The regression weights of the predictor scores for Sample B (the same scores and weights used in the interim battery) were also applied to Sample C responses. The scores of sample C recruiters were calculated and validated against the pass/fail criterion and yielded a point biserial correlation of .21 ( $p < .05$ ).

The same procedures were followed in validating the predictor scores against a second criterion--supervisor field ratings. These ratings were obtained from supervisors one year after each recruiter was placed on duty. Results showed that no combination of variables of the entire set significantly predicted the field ratings at the .05 level.

Two other variables outside the predictor battery were also investigated for predicting recruiter success. First, the recruiter's primary Air Force Specialty Code (AFSC), or the service specialty in which a recruiter had previously worked, was investigated because recruiter training personnel believed that those specialties requiring more interpersonal contact would produce better recruiters than those specialties that require little contact. The primary AFSC variable was divided into four service specialty variables reflecting various degrees of required interpersonal contact. Additionally, 12 interaction variables were generated by multiplying each of the four primary service variables by aptitude, surgency, and dependency scores, thus creating 12 hybrid individual differences/AFSC variables. These 16 variables were then correlated with the field-rating criterion measure for a sample of 859 recruiter course graduates. Only one of the 16 correlated as highly as .07 ( $p < .05$ ). The authors therefore discarded the primary AFSC as a predictor of recruiter success.

A second variable regarded as a possible predictor of field recruiter performance was recruiter school advisor ratings. These ratings, collected while the recruiters were in recruiter school, consisted of a single 7-point scale ranging from outstanding to unsatisfactory. The advisor ratings were related to field ratings, yielding a correlation of .19 ( $p < .05$ ).

The results of this study indicate that a battery consisting of the Recruiter Language Test, Community Information Inventory, and the FCSRI-A (Surgency) can be used to predict recruiter school performance, but that neither this battery nor recruiter school advisor ratings can be used to predict field recruiter performance. The authors attribute the failure to predict field recruiter performance to the supervisory performance rating criterion measure. They suggest that this kind of rating is inherently contaminated by several rater errors such as leniency. Another difficulty may be that the battery the authors developed is simply more appropriate for predicting recruiter school success than field recruiter performance. At any rate, Massey and Mullins suggested that a more reliable and valid criterion measure should be developed to provide a fair assessment of the utility of any paper-and-pencil measures purporting to predict recruiter performance.

Borman (1982). All the studies discussed thus far employed paper-and-pencil measures to predict recruiter performance or effectiveness. By contrast, Borman (1982) developed assessment center exercises targeted

toward predicting Army recruiter performance in training. The assessment center exercises were designed to tap 14 personal characteristics (e.g., Sociability, Personal Impact) that experts on the Army recruiter job had rated as important for performing recruiter tasks [Borman, Toquam, & Rosse (1976), provides a summary of the initial job analysis]. The six exercises developed by Borman were:

1. Structured interview. Assessors ask a series of questions targeted at the subject's level of achievement motivation, potential for being a "self-starter", and commitment to the Army.
2. Cold calls. Subject has an opportunity to learn a little about three prospects (e.g., Joe Hill and Jack Brown) and must phone each of them for the purpose of getting them to come into the office. Assessor role players have well-defined characters (prospects) to portray.
3. Interviews. Two of the three cold-call prospects agree to come in for an interview. The subject's job is to follow up on what was learned in the cold-call conversations and to begin promoting Army enlistment to these people. A third "walk-in" prospect also appears for an interview with the subject.
4. Interview with concerned parent. Subject is asked to prepare for and conduct an interview with the father of one of the prospects he or she interviewed previously.
5. 5-minute speech about the Army. Subject prepares a short talk about an Army career that he or she delivers to the rest of the group and to the assessors.
6. In-basket. Subject is given an in-basket filled with notes, phone messages, and letters on which he or she must take some action.

Borman conducted a two-day course to train 16 experienced and successful Army recruiters as assessors. Fifty-seven soldiers entering the U.S. Army recruiter school with no previous recruiting experience were assessed in the evaluation program. Twelve assessors were assigned in pairs to subjects in the structured interview and five role-playing exercises. Four additional assessors, working in pairs, provided "first impression" ratings, or evaluations of subjects made after only brief exposure to them during a 15-minute introductory session at the beginning of the assessment day. The author made physical attractiveness and likeability ratings of each subject on the basis of interaction with them in the introductory meeting and during exercise instruction sessions with them before each exercise.

Assignments of assessors to subjects and exercises were made such that (1) assessor pairs never worked together more than one day during the ten-day program to avoid possible rater effects on interrater agreement results; (2) assessors were assigned to the main group of 12 for seven or eight of the ten days and to the group of four for the other days; and (3) each pair viewed each subject in one and only one of the six exercises during the day. This latter feature provided the opportunity to evaluate the relative contribution of each exercise to validity.

For each exercise, assessors rated each trainee on the personal characteristics relevant for that exercise and provided an overall performance rating for each exercise. At the end of each day, the assessors discussed subjects' performance and reached a consensus on their ratings. Thus, the assessment center design provided five types of predictor measures:

1. First impression, physical attractiveness, and likeability ratings;
2. Ratings on a structured interview;
3. Ratings of subjects' performance on individual exercises;
4. Consensus ratings of subjects' performance; and,
5. A statistical composite (unit weighted) of the assessment ratings.

Additionally, Borman administered the SAB (Borman et al., 1980) and a personality inventory (Brown et al., 1975) as predictor measures.

Three measures obtained during two phases of training were employed as criteria. A Phase 1 composite of scores on three objective tests comprised one criterion. The tests measured mastery of prospecting and selling techniques taught in Phase 1. An intraclass correlation of .94, indexing the composite's reliability, suggested that the Phase 1 criterion measures were homogeneous.

In Phase 2, students practiced the telephoning and interviewing techniques they had learned in Phase 1. In small groups and with the help of a staff instructor, they practiced their skills in hypothetical recruiting situations. Between the fifth and the eighth day after students had entered Phase 2, instructors rated these students on three performance dimensions. In general, individuals who were performing best in Phase 2 were released earliest ( $r = -.52$  between a composite of the Phase 2 instructor ratings and number of days to complete Phase 2). The Phase 1 composite correlated .72 and -.33 respectively with instructor evaluations in Phase 2 and time to complete Phase 2. It should be noted that instructors in the training course had no contact with the assessment program or the assessors. Thus, ratings they made of trainee performance during training were completely independent of assessment ratings.

Interrater agreement between assessors on each personal characteristic for each of the six exercises ranged from .44 to .92 with a median of .76. Ratings of overall performance were more reliable (median  $N = .84$ ) than ratings on individual personal characteristics. Validity coefficients between the predictors and three criteria are presented in Table 6.

As shown, first impression, physical attractiveness, and likeability ratings showed nonsignificant validities, as did the structured interview, the SAB, and the personality test. On the other hand, ratings on all five role-playing exercises yielded correlations of .32 ( $p < .05$ ) or greater with the Phase 1 composite. Correlations were not as high for the Phase 2 criteria, performance ratings and time-to-completion.

TABLE 6  
Correlations Between Assessment Measures and Training Criteria  
(Borman, 1982)

Variable	Phase 1	Phase 2	
	Composite performance	Composite performance	Time to complete
Special Assignment Battery (SAB) test keys (Navy)			
Human relations	-06	15	-27
Selling	02	17	-12
Organizing	-01	07	-15
Overall performance	-09	15	-09
Army personality test			
Experimental key	04	05	-21
Assessment measures			
First impression	23	-18	-02
Physical attractiveness	-06	-20	06
Likeability	06	04	14
Structured interview	07	26	00
Assessment ratings			
Cold calls	33*	-03	-26
Interviews	34*	15	-28
Interview with parent	41**	24	-29*
In-basket	33*	36*	-11
Speech	32*	27	-42**
Consensus rating	38**	27	-24
Unit-weighted composite			
Ratings on relevant personal characteristics (all exercises)	48**	35*	-33*
Corrected for attenuation	53**	50**	-49**

Note: For Phase 1, N = 50; For Phase 2, N = 46.

Decimal points are omitted.

\*  $p < .05$  (two-tailed). \*\*  $p < .01$  (two-tailed).

The unit-weighted composite of the assessment (i.e., the sum of the scores on each personal characteristic pooled across the six exercises) correlated significantly with each of the three criteria (.48 with the Phase 1 composite; .35 with the Phase 2 composite; and -.33 with time-to-complete Phase 2). Borman noted that these validities were conservative because attrition occurred during the study. Four trainees with relatively low assessment scores were eliminated from training because of inadequate Phase 1 performance. Seven persons who received either very high or very low assessment ratings dropped out before training. Thus, the range of assessment ratings was restricted. Corrected for restriction of range, validity coefficients for the unit-weighted composite were .53, .50, and -.49, respectively against the three criterion measures.

Borman, Eaton, Bryan, and Rosse (1983). This study further explored the validity of the assessment exercises developed by Borman (1982) for predicting success in recruiter training. A subset of the exercises employed previously by Borman was administered to 450 sergeants entering the U.S. Army's recruiter school. These exercises included:

1. "Cold call" with Joe Hill, a prospective recruit unwilling to consider Army enlistment;
2. "Cold call" with Jack Brown, a prospect willing to consider Army recruitment;
3. 5-minute speech about the Army; and,
4. In-basket exercise.

Thirteen experienced recruiters received assessor training and served as assessors. For each exercise, assessors evaluated assessees on the personal characteristics relevant for that exercise and on an overall performance dimension.

Similar to the 1982 study, three training performance criteria served as elements of the criterion composite. One element was a 100-item objective test measuring mastery of prospecting and selling techniques taught in the first phase of recruiter school. The second and third elements were instructor ratings of telephone and interview selling techniques taught in a second phase of recruiter training. The telephone and interview performance ratings, made by different instructors, correlated .28, and each set of ratings correlated .24 with the test scores. A unit-weighted composite of these three criteria was used as the primary criterion in this research.

Correlations between the assessment ratings and training performance scores were computed to investigate the validities of the various exercises as predictors of training performance criteria. The validities for the four individual exercises ranged from .10 to .26, with a median of .19.

The correlation between the unit weighted composite of the exercise ratings and the composite criterion was .32. Although these validities were not impressively high, they were significantly different from zero ( $p < .01$ ).

Borman et al. (1983) suggest that the lower validities obtained in this study compared to previous work (Borman, 1982) may have been the result of

changes in the performance criterion measures taken at the recruiter school. Another possibility is that the assessors may have been distracted from predicting performance by their additional responsibilities of providing performance feedback after each exercise.

In sum, the validity of both paper-and-pencil measures and assessment scores for predicting recruiter success in training has been investigated. Massey and Mullins (1966) found that paper-and-pencil measures of personality characteristics and verbal ability successfully predict completion of recruiter training (pass/fail). Paper-and-pencil measures of non-cognitive variables did not, however, correlate significantly with objective training test scores, instructor ratings, or time to complete training criteria employed by Borman (1982). Borman also designed assessment center exercises to tap non-cognitive and cognitive personal characteristics and employed assessment scores as predictors. Because both non-cognitive and cognitive characteristics are tapped by such exercises, no conclusion regarding the relative merit of either type of variable as a predictor can be drawn. Of particular importance though, this research provides evidence for the utility of the assessment approach for predicting both objective and subjective training criteria.

## Measuring and Predicting Performance in Military Recruiter and Civilian Sales Jobs: An Integration

In previous sections we described research conducted to identify the person characteristics and environmental factors associated with military recruiter success. At a broader level, however, the military recruiter job is one among many sales jobs, and within the civilian sales arena numerous studies have been conducted to delineate variables related to sales effectiveness. We, therefore, prepared a review of civilian sales selection research (Kanfer & Borman, 1986) and analyzed the similarities and differences among factors underlying military recruiter and civilian sales selection research.

Three major themes emerged from our comparison of these two lines of research (i.e., military recruiter and civilian sales selection). First, although recruiting is essentially selling, differences in the type of "product" sold and clientele served result in some differences between the military recruiter job and sales jobs in the civilian sector. For instance, recruiters are selling a lifestyle or a career rather than material goods or services. Such differences between military and civilian sales jobs lead to somewhat different behavioral requirements in the two arenas and, in turn, can moderate the validity of specific predictors against performance criteria. The usefulness of civilian sales selection research for drawing detailed conclusions about the variables likely to predict military recruiter performance is, thus, somewhat limited.

Second, the major points regarding the quality of various criteria discussed throughout previous chapters of this review apply in both civilian and military sales arenas. That is, environmental factors may restrict or enhance the range of possible outcomes of performance, and thus contaminate objective effectiveness measures, and objective indices may not tap all aspects of performance. Subjective measures (such as supervisor ratings) can provide multidimensional evaluations, are probably not restricted by environmental factors, and thus offer a distinct advantage in selection research. The quality of subjective measures is, however, influenced by the source of the rating and various rating errors and biases. Cumulating research across civilian and military jobs is, thus, useful for identifying potential pitfalls and problems associated with various criteria, understanding the interrelationships among criteria, and developing a strategy for obtaining criterion data of high quality.

Finally, the major finding of our comparison is that some of the personal attributes that lead to success in recruiting are, in general, similar to those requisite to effective civilian sales performance. In particular, personality measures tapping dominance or forcefulness appear to be associated with both recruiter and sales success, and in contrast, aptitude measures have little support as predictors of success in either arena. Vocational interest variables frequently found to be good predictors of military recruiter success have not been researched sufficiently against civilian sales performance. Similarly, little research has investigated the validity of skill level variables (e.g., objective assessment of interpersonal skills) for predicting military recruiter job

performance. Because such variables have yielded reasonably good validities in sales selection research, they may also be useful predictors of recruiter job performance.

The following three sections summarize recruiter selection research described previously in this review, integrating relevant research from the civilian sector. In particular, the first section discusses similarities and differences between civilian sales and military recruiter jobs, and the second section focuses on criterion-related research to explicate future research issues. The third section summarizes validation research for specific predictor variables in both military and civilian sales settings.

#### The Performance Domain

How similar are the performance requirements of the military recruiter job to those of civilian sales jobs? In the broadest sense, all sales jobs involve behaviors on the part of the salesperson to sell and promote a product to a client. Differences in the behaviors requisite to effective performance in sales jobs, however, may emerge from differences in the types of products sold or the types of clientele served.

Sales jobs are extremely diverse in terms of the types of products sold (e.g., life insurance, computers, services) and the clientele served (e.g., individual customers, businesses). How does this diversity impact on predictor-criterion relationships? In a recent meta-analysis of civilian sales research, Churchill et al. (1985) categorized products as consumer goods, industrial goods, or services, and found product type to moderate the relationship between predictors and criteria. More specifically, product type had the greatest moderating effect on the relationship between non-cognitive variables (e.g., motivation, personality) and sales success measures. The authors also investigated moderating effects of customer type (independent of product type) by classifying customers as either individual consumers or institutions. The results suggested that customer type moderates relationships between non-cognitive predictors and measures of sales success.

Two aspects of this work warrant further attention. First, the product and customer types categories employed by Churchill et al. (1985) are quite broad. Such breadth could attenuate the moderating effects of these variables by increasing the variance within moderator type. Second, the product type and customer type variables employed in this work are tapping redundant information. For example, most industrial goods are probably sold to institutions. Perhaps categorizing products more specifically (e.g., vehicles, appliances, insurance/benefits, etc.) would further distinguish moderating effects of these variables. Of particular importance, though, this research suggests that product and client types moderate predictor-criterion relationships even when these factors are categorized broadly.

With respect to product and client type the military recruiter job is quite different from most sales jobs in the civilian sector. Indeed, military recruiters are selling a concept, a career, a lifestyle, or an education rather than a "product." On the "client" side, military recruiters are faced with selling/projecting an image of the armed services to the community, as well as selling military careers to young adults individually.

Because products and clients define many of the knowledges required by the sales job, they impact on behaviors in which job knowledge is exhibited. Additionally, specific products or clients may necessitate a unique set of job behaviors. To explicate these ideas, task and performance categories developed through job analyses of Navy/Marine Corps (Borman et al., 1976) and Army (Borman, Toquam, & Rosse, 1976b) recruiter jobs are provided in Table 7. The four Army recruiter task dimensions resulted from an empirical analysis of task responses (Borman et al., 1976b) and actually subsume 26 more specific task categories. The eight Navy/Marine Corps recruiter performance dimensions were developed by the critical incident or behavior analysis methodology and were described in an earlier section of this review. Also included in Table 7, are recent results of Army recruiter performance rating scale development research (Borman, Russell, & Skilling, 1986). Consistently across these studies we have found that military recruiter performance involves a number of selling behaviors such as prospecting, developing rapport with clients, and making sales presentations, and a number of "other" behaviors. The "other" behaviors focus on community relations, organizational and co-worker commitment, planning and organizing, and exhibiting knowledge of the programs available.

For comparison purposes, two job analyses of civilian sales positions (Behrman & Perreault, 1982; Kirchner & Dunnette, 1957) are also included in Table 7. The individual customer-oriented sales activities appear similar across these jobs. One difference between performance requirements of the two (i.e., recruiter and civilian sales) is the emphasis on community-oriented sales behaviors in the military recruiter job. More specifically, civilian sales performance, does not appear to emphasize publicizing the organization and maintaining rapport with persons in the community. Perhaps the nature of recruiter "products" (i.e., lifestyle, career) and clients (i.e., young adults in the community) requires community-oriented selling behaviors as well as individual customer-oriented selling activities.

In summary, civilian sales jobs are diverse in terms of "products or clients," and unlike most civilian sales jobs, the military recruiter job does not involve selling a product or a service. Differences in product and client type probably lead to unique performance requirements across civilian sales jobs (as well as between military recruiter and civilian sales jobs) and, in turn, impact on predictor-criterion relationships.

#### Measuring Performance and Effectiveness

In addition to variations in performance requirements across sales positions, methods of measuring performance or effectiveness may contribute variation in predictor-criterion relationships. Performance is typically measured using subjective ratings (e.g., peer, self, and/or supervisory ratings). Effectiveness measures, on the other hand, are objective indices of a "bottom-line" outcome of performance (e.g., sales volume, recruiter production). In previous sections of this review, we discussed issues regarding objective and subjective criteria and military recruiter research that addresses these issues. The following paragraphs

Table 7  
RECRUITER AND CIVILIAN SALES JOB ANALYSES

Type of Category	Military Recruiter Job Borman, Mough & Dunnette	Military Recruiter Job Borman, Toquam, & Rosse	Civilian Sales Jobs Borman, Russell, & Skilling	Civilian Sales Jobs Kirchner & Dunnette	Civilian Sales Jobs Behrman & Perreault
<u>Individual-Selling</u>					
Locating/prospecting clients	Locating and contacting qualified prospects	Prospecting activities	Locating and contacting qualified prospects	Calling on new accounts; persisting on tough accounts	
Building rapport	Gaining and maintaining rapport		Gaining and maintaining rapport	Following up; carrying out promises; helping customers	Developing and maintaining customer goodwill
Identifying client needs and selling skills	Obtaining information from prospects and making good Navy-person fits; selling skills	Selling activities	Obtaining information from prospects and determining their needs and interests; selling skills	Knowing customer requirements; initiating new sales ideas; pointing out other uses for products; using new sales techniques/ methods; showing a non-passive attitude	Giving high quality sales presentations and working well with customers; achieving quantity and quality sales
<u>Other</u>					
Exhibiting technical knowledge	Communicating knowledgeable and accurate information		Communicating knowledgeable and accurate information	Communicating necessary information; communicating truthful information	Developing and using technical knowledge
Organizational commitment	Supporting other recruiters and the command		Supporting other recruiters and USAREC	Defending company policy; preventing price cutting	Providing information to the company and following company policy; controlling unnecessary expenses; working well with other personnel in the firm
Planning and organizing	Organizing skills	Administrative activities	Organizing skills	Planning ahead	
Community relations	Establishing and maintaining good relationships in the community	Publicizing Army	Establishing and maintaining good relationships in the community		

summarize major findings regarding performance measurement in the recruiting setting, incorporating information from research in the civilian sector, and focusing on implications regarding future research.

### Objective Measures

Although "bottom-line" effectiveness measures such as sales volume and recruiter production are objective criteria, their usefulness in selection research is limited by two problems. First, as a single index, an effectiveness measure may not tap all aspects of the performance domain. Second, effectiveness measures may be affected by environmental factors not within the individual's control. As a consequence, such measures may not accurately reflect variation in sales effectiveness due to individual differences in behavior and/or performance.

What does an effectiveness criterion measure? Many objective measures have been employed in civilian sales research (e.g., sales volume ranking, dollar volume of yearly sales, earnings, commissions, average number of sales), and observed correlations among different measures are sometimes high and sometimes low (cf. Kanfer & Borman, 1986). Low correlations could indicate either unreliability (if the measures are indexing the same construct) or, if the measures are expected to tap different aspects of sales success, actual differences in the constructs measured. Are different objective measures indexing different aspects of sales success? What relationships exist between objective indices and subjective measures? A small number of studies have examined such relationships. Rush (1953) obtained correlations between several objective and subjective sales criteria and factor analyzed the correlation matrix. Borman et al. (1979) factor analyzed correlations between Navy recruiter rating dimensions, resulting in an interpretable three-factor solution. Factor scores on each rating factor (Selling, Human relations, and Organizing Skills) were then correlated with a summary, objective production index, and the latter index related most highly with the Selling factor, as hypothesized. However, further research focusing on relationships between different objective measures and between objective and subjective criteria is needed to clarify factors that underlie sales performance.

What approaches might enhance the quality of objective measures? A number of researchers in both the recruiter (e.g., Brown et al. 1975; Larriva, 1975) and civilian sales (e.g., Miner, 1962) arenas have developed adjusted indices to reduce the contaminating, non-performance related variance in effectiveness criteria. It is, however, difficult to assess the validity of an adjusted measure because it is typically not known whether the adjustment is reducing error or "true score" variance (or both). The rationale behind the type of adjustment is, therefore, very important to the construct and content validity of the index.

One common adjustment made on objective effectiveness indices is to adjust for mean differences in production between recruiting territories. Referring to the point made above, unfortunately, to the extent there are meaningful, "true" differences in individuals' performance between territories, such an adjustment removes actual variance in effectiveness. Research suggests that adjusting for differences between territories does remove some contaminating variance due to the environment; however, it is also possible that environmental factors contribute contaminating var-

iance within territories (Borman et al., 1982). Another way to adjust objective measures is to identify environmental factors beyond the salesperson's or recruiter's control and then partial out of the production index variance due to these variables.

Regarding environmental factors, recall that objective effectiveness indices are measures of an organizational outcome of performance. Environmental factors are characteristics of the sales territory (or recruiting area) that restrict or enhance the range of possible outcomes (effectiveness). Research has been conducted to delineate environmental variables (e.g., Cravens & Woodruff, 1973; Borman et al., 1982). Examples from the case of military recruiting are: (1) unemployment rate, (2) military/civilian job pay ratio, and (3) propensity to enlist. To the extent different territories are very different with regard to these environmental variables, production in individual territories might be either enhanced or inhibited. Unfortunately, researchers (e.g., Arima, 1977; Brown et al., 1975) have sometimes failed to distinguish between the territory itself as a variable and the more specific variables describing the recruiting/selling setting, confusing the issue. Borman et al. (1982) point out that adjusting for differences in territory, by definition, accounts for all between-territory environmental differences. In contrast, the environmental variables considered in a study can only approach accounting for all between-territory variance. In this sense, simply adjusting for mean territory differences is better than the partialling procedure used with environmental variables included in the study. The only problem with this method of adjustment is that, as mentioned, it may remove true performance as well as error (environmental) variance from production scores. More research is needed to examine these issues, but the researcher using production data should be aware of these possible pitfalls and problems with different adjustment strategies.

### Subjective Measures

Subjective ratings of sales performance have also been employed as criteria in both recruiter and civilian sales selection research. Because such ratings can provide multidimensional evaluations of performance and are not restricted by environmental factors, they offer a distinct advantage in selection research. The quality of subjective ratings is, however, influenced by the source of the rating and various rating errors and biases.

How does the source of the rating affect rating quality? Summarizing the advantages and disadvantages of various rating sources, Kanfer and Borman (1986) state:

"Subjective ratings may be obtained from managerial personnel, peers, independent observers, and/or self-report. Each rating source has its advantages and disadvantages. Supervisory ratings typically come from persons with good experience in rating performance, but supervisors may be less familiar than other sources with ratees' day-to-day activities. Peers are usually very familiar with incumbents' actual behavior related to performance, but they often have less experience (than supervisors) in completing ratings and are more likely to provide lenient (higher than deserved) evaluations. Observers may be less prone to common rating errors (e.g., halo and leniency) and be more

objective, but they lack the longer term knowledge of ratees' performance over time. Finally, self-ratings obviously come from a source that should know well the true criterion performance levels, but these ratings tend to be inflated (i.e., lenient)" (p. 11).

How might the quality of subjective ratings be enhanced? Ratings are typically made along a number of dimensions. When such dimensions (e.g., leadership) are not defined by the rating scale, individual raters may adopt different definitions and standards for performance. Thus, theoretically, rating scales that specifically define the behavioral performance requirements or standards within a dimension and describe the behaviors associated with various levels of performance should lead to less subjectivity in ratings. Format comparison research has not demonstrated a clear superiority of any one specific type of approach to designing formats (e.g., Landy & Farr, 1980). However, considerable research in applied settings where "for-research-only" peer or supervisor ratings have been gathered using behaviorally oriented scales have resulted in reasonably reliable ratings, demonstrating some degree of construct validity, as well (Borman et al., 1979, 1980; Bownas, Rosse, & Dunnette, 1977; Peterson & Houston, 1980; Pulakos & Borman, 1986). Interrater reliabilities typically reach .60-.70, factor analysis results are often interpretable, such as with the Selling, Human Relations, and Organizing Skills factors identified in the Navy recruiter research, and correlations between rating factors and other criteria as well as with predictors have been conceptually appropriate (e.g., Selling Skills correlating more highly with production than the other two rating factors and the personality predictor scale, Order, relating higher with Organizing Skills than with the other rating factors in the same study). To clarify, we are not asserting that ratings have no or even very few problems. Difficulties with rating errors should be acknowledged (e.g., leniency, halo, and the like). However, steps can and have been taken to improve the reliability and validity of performance ratings.

These positive results with ratings seem to be a function of the combination of behavior-based scales and training (e.g., Pulakos & Borman, 1985) to orient raters to the format and to reduce certain rating errors. While no experimental results are available to conclude that this format and training "treatment" is necessarily most effective, it is reassuring that criterion measurement using this combination has proven to be reasonably successful in actual practice.

We have discussed several factors contributing to variance in measures of performance. First, product and client types probably contribute some true differences in the behaviors required for effective performance across sales positions. Second, error in subjective ratings can contribute to error in these performance criteria. However, steps can be taken to reduce some of this error. Third, environmental factors may restrict or enhance the range of possible outcomes of performance and, thus, contaminate objective effectiveness measures, and, finally, objective indices may not tap all aspects of performance. All of these factors could attenuate observed predictor-criterion relationships across sales and recruiter positions.

## Personal Attributes Important for Effective Performance

Individual differences variables used to predict the performance of sales personnel and military recruiters can be classified into five categories: biographical, personality, interest, skill-related, and aptitude variables. The paragraphs below briefly summarize military recruiter and civilian sales selection research with respect to these predictors. More information on military recruiter selection studies is provided in previous chapters of this review, and civilian sales selection research is discussed in greater detail by Kanfer and Borman (1986).

### Biographical Variables

Biographical variables represent a diversity of items such as age, prior selling experience, marital status, work habits, and education level. Unlike measures of personality, interest, and aptitude in which multiple items tap a construct, a single biographical item is often examined as a predictor.

Recruiter research. Biographical variables have met with some success as predictors of military recruiter performance. Individual variables such as educational background, prior service classification, marital status (e.g., Brogden & Taylor, 1949; Krug, 1972), work habits, and styles of handling finances and debts (Brown et al., 1975) have sometimes produced significant correlations with performance; however, further research is needed to replicate these results. The Special Assignment Battery (SAB) includes 136 biographical items. Of these only seven have consistently yielded good relationships with recruiter performance (Borman et al., 1979; 1980; 1983). These items pertain to planning ahead, involvement in and attitude toward work, past recognition from teachers, number of towns lived in, and attentiveness to others. Interestingly, they tend to be related to performance in organizing and less associated with other aspects of recruiter performance.

Civilian sales research. Aggregating across types of sales jobs and types of biographical variables, Churchill et al. (1985) obtained a mean corrected correlation of .29 between biographical factors and performance criteria in a recent meta-analysis of civilian sales selection research. The authors also found that the type of product moderated biographical predictor-criterion relationships. More specifically, corrected mean correlations were .43, .29, and .21 for jobs involving service sales, consumer goods, and industrial goods respectively. Kanfer and Borman (1986), suggested that these relationships might be attenuated by aggregation across biographical items that may tap different underlying constructs and examined results for individual biographical predictors. Age, family history and status, education level, previous selling experience, and other biographical variables, however, did not appear to consistently predict sales performance.

Summary. Recruiter and civilian sales research suggests that biographical predictors may be useful predictors of sales performance. Research to date, however, has not delineated which biographical variables are predictive of sales success in certain types of sales jobs. Findings for individual variables such as education level and family history are somewhat inconsistent or unreplicated within as well as across these two

lines of research.

Three conceptual issues preclude broad generalizations about the utility of biographical variables, in general, as predictors of sales success. First, research suggests that biographical predictor-criterion relationships are moderated by the type of product. Aggregating across types of sales jobs is thus likely to obscure actual predictor-criterion relationships. Second, as discussed by Kanfer and Borman (1986) the broad delineation of personal history variables implies that multiple factors may underlie biographical predictors (e.g., perhaps, biographical variables are indexing personality and interest constructs). Aggregating across biographical variables tapping different underlying factors would attenuate effects of these predictors. The third problem relates to the fact that correlations between individual biodata predictor items and performance are often reported. Because single items are expected to be less potent psychometrically than a scale of items, low correlations do not preclude the predictive validity of biodata and may instead indicate a need for multiple-item biodata scales. Further research is therefore needed to explicate relationships between biographical and other variables and between biographical variables and performance before biographical items can be dismissed as good predictors of sales success.

### Personality Variables

Personality measures are designed to index traits--stable non-cognitive attributes--that influence behavior. In sales and recruiter research the validity of traits such as empathy, self-esteem, and dominance for predicting performance has been investigated.

Recruiter research. A number of personality variables typically produce significant correlations with recruiter performance or production criteria (e.g., Atwater et al., 1986; Borman et al., 1979, 1980, 1983; Krug, 1972; Larriva, 1975). Personality inventories most often used in these studies are the Special Assignment Battery and the 16PF. Items tapping characteristics such as dominance, aggressiveness, self-confidence, and spontaneity appear to be most consistently associated with recruiter performance. Surgency, cooperativeness, and orderliness have shown relationships with training criteria, but not with supervisors' ratings of job performance (Massey & Mullins, 1966). Likewise, empathy, achievement motivation, and sociability were not related to performance in a study by Brown et al. (1975).

Civilian sales research. Aggregating across personality constructs and across types of sales jobs, Churchill et al. (1985) obtained a mean corrected correlation of .26 between personality and motivation variables and sales performance. When individual personality traits are examined separately, certain results across studies appear quite consistent. For example, personality scales tapping dominance and forcefulness are often associated with civilian sales performance (e.g., Dunnette & Kirchner, 1960; Harrell, 1960; Lamont & Lundstrom, 1977; Rodgers, 1956). Results for self-esteem variables have been positive (Bagozzi, 1980; Ghiselli, 1969); however, nonsignificant findings are also present (Gable, Mattheiss, & Muczyk, 1984). Empathy measures (e.g., Cotham, 1969; Lamont & Lundstrom, 1977) and achievement motivation variables (e.g., Bagozzi, 1980; Ghiselli, 1969; Matteson, Ivancevich, and Smith, 1984) have both provided mixed re-

sults. Items designed to tap supervisory ability were significant predictors in one study (Ghiselli, 1969) and initiative produced positive results in another (Gable et al., 1984). Social poise as indexed by the Minnesota Multiphasic Personality Inventory (MMPI) K scale has also yielded a positive relationship with performance (Ruch & Ruch, 1967).

Summary. Recruiter and civilian salesperson performance both appear to be related to personality attributes. More specifically, variables indexing dominance and forcefulness appear to be related to both recruiter and salesperson performance. Self-esteem and confidence have yielded positive relationships in both recruiter and civilian sales contexts. Empathy and achievement motivation variables, on the other hand, do not appear to be highly associated with performance in either arena.

Aggregating findings across personality traits and across types of sales jobs can have complex and attenuating effects on predictor-criterion relationships. For instance, product and client type were found to be strong moderators of personality/motivation predictor-criterion relationships in the Churchill et al. (1985) meta-analysis. Indeed, it is possible that personality variables are, in general, more or less predictive depending upon the type of sales job being studied. It is also possible that specific personality traits do underlie performance in most sales jobs while other traits are related to performance in specific jobs or are not related to sales performance at all. Combining personality variables does not explicate these relationships.

### Vocational Interest Variables

Recruiter research. Vocational interest variables have proven to be good predictors of both recruiter performance (cf. Abrahams et al., 1973; Borman et al., 1979; 1980; 1983; Graf & Bower, 1976; Wollack & Kipnis, 1960) and recruiter production (Atwater et al., 1986). Many of these studies have involved development of empirical keys of valid items, and some have not discussed the type of items characterizing the key. Items referring to interest in extroverted, dominant, or competitive activities and interest in politics and law are examples of those related to criteria when types of valid items are reported (cf. Atwater et al., 1986; Borman et al., 1979, 1980, 1983; Wollack & Kipnis, 1960). The Strong-Campbell Interest Inventory and the Kuder Preference Record are the interest inventories most often employed in recruiter selection research.

Civilian sales research. As discussed by Kanfer and Borman (1986) civilian sales research has focused on comparisons of response patterns across types of sales jobs rather than using interest variables to predict performance. The validation research that does exist suggests that the type of sales job may influence the predictive validity of vocational interest measures (Dunnette & Kirchner, 1960).

Summary. While there is a relative dearth of research exploring predictive validity of vocational interest measures against civilian sales performance, interest variables have been investigated frequently and with success in military recruiter research.

## Skill Level Variables

Skill level measures are designed to assess the individual's current level of functioning with respect to specific job activities and are thus often developed specifically for the sample under investigation.

Recruiter research. The only recruiter selection study investigating the utility of skill level variables was conducted by Brown et al., 1975. Recruiters were asked to simulate a presentation of the benefits of Army life to a prospective enlistee. Presentations were recorded and scored. The scores differentiated significantly between two (high and low) criterion groups. Skill level variables were evaluated against training criteria in two studies by Borman and his colleagues (Borman, 1982; Borman et al. 1983). In both studies, performance on four assessment center exercises--cold call with prospect, follow-up interview with prospect, five minute speech about the Army, and in-basket--was significantly correlated with training criteria.

Civilian sales research. Interpersonal skills such as the ability to identify verbal and non-verbal cues (Weitz, 1979), to maintain eye contact, and to produce a good impression (Pace, 1962) have shown some association with sales performance. Churchill et al. (1985) reported a corrected mean correlation of .32 between skill level predictors and performance. The authors also found these predictors to be less influenced by the product and client type than are personality and biographical variables.

Summary. The fact that little recruiter selection research has investigated the utility of skill level predictors limits comparison of military recruiter and civilian sales arenas. The Borman et al. (1982, 1983) research does suggest that assessment centers are valid predictors of training criteria; however, the utility of these variables for predicting recruiter performance on-the-job has not been explored. Research in the civilian sector does suggest that interpersonal skill variables may have utility for predicting job performance.

## Aptitude Variables

Recruiter research. General aptitude and verbal ability measures have yielded little or no relationship with recruiter performance criteria. For example, neither of two verbal ability tests employed by Massey and Mullins (1966) predicted performance ratings. Similarly, only one out of 24 verbal ability-performance relationships was significant in a study by Wollack and Kipnis (1960). Aptitude measures of general intelligence, arithmetic reasoning, and mechanical comprehension have also shown no relationship with performance criteria (cf. Brown et al., 1975; Wollack & Kipnis, 1960).

Civilian sales research. Aptitude measures have met with little to mixed success as predictors of sales criteria. For instance, Harrell (1960) and Bray and Campbell (1968) both found general ability measures to be correlated with sales performance. Miner (1962), however found that the only Wechsler Adult Intelligence Scale correlating with an objective criterion was the arithmetic subtest. Rush (1953) found that business arithmetic and college grades were positively correlated with an objective criterion. Verbal intelligence, on the other hand has correlated negatively with sales success (Bagozzi, 1980). Of the various person attribute pre-

dictors investigated in the Churchill et al. (1985) meta-analysis, aptitude measures produced the lowest predictor-criterion correlation (attenuation corrected  $r = .19$ ), and no moderating effects due to product or client types were observed.

Summary. In general, aptitude measures have shown little or no utility for predicting recruiter or civilian sales performance.

#### Summary: Predictor-Criterion Relationships

Our review of military recruiter and civilian salesperson selection research suggests that there are two major similarities between observed predictor-criterion relationships. First, both recruiter and salesperson performance appear to be related to personality attributes. In particular, personality variables indexing dominance and forcefulness are likely to be associated with recruiter and salesperson performance. Other personality variables appear to yield less consistent relationships. It is possible that the validity of specific personality variables is moderated by the type of sales job. To date, research does not clarify this issue. The second major similarity is the lack of support for aptitude measures as predictors of either sales or recruiting performance.

## References

Abrahams, N. M., Neumann, I., & Rimland, B. (1973). Preliminary validation of an interest inventory for selection of Navy recruiters (NPTRL Research Memorandum SRM-73-3). San Diego, CA: Navy Personnel and Training Research Laboratory.

Arima, J. K. (1977). Determinants and a measure of Navy recruiter effectiveness. Prepared for the Navy Personnel Research and Development Center, San Diego.

Atwater, D. C., Abrahams, N. M., & Trent, T. T. (1986). Validation of the Marine Corps Special Assignment Battery (SAB). San Diego: Navy Research and Development Center.

Baehr, M. E., & Williams, G. B. (1967). Underlying dimensions of personal background data and their relationship to occupational classification. Journal of Applied Psychology, 51, 481-490.

Baehr, M. E., & Williams, G. B. (1968). Prediction of sales success from factorially determined dimensions of personal background data. Journal of Applied Psychology, 52, 98-103.

Bagozzi, R. P. (1978). Salesforce performance and satisfaction as a function of individual difference, interpersonal, and situational factors. Journal of Marketing Research, 15, 517-531.

Bagozzi, R. P. (1980). Performance and satisfaction in an industrial sales force: An examination of their antecedents and simultaneity. Journal of Marketing, 44, 65-77.

Baier, D. E., & Dugan, R. D. (1957). Factors in sales success. Journal of Applied Psychology, 41, 37-40.

Baker, E. M., & Schuck, J. R. (1975). Theoretical note: Use of signal detection theory to clarify problems of evaluating performance in industry. Organizational Behavior and Human Performance, 13, 307-317.

Behrman, D. N., & Perreault, W. D., Jr. (1982). Measuring the performance of industrial salespersons. Journal of Business Research, 10, 355-370.

Bennett, J. T., & Haber, S. E. (1973). Selection, deployment, and evaluation of Marine Corps recruiters. Washington, DC: The George Washington University, School of Engineering and Applied Science, Institute of Management Science and Engineering, Project NR347020, Office of Naval Research.

Borman, W. C. (1982). Validity of a behavioral assessment for predicting military recruiter performance. Journal of Applied Psychology, 67(1).

Borman, W. C., Eaton, N. K., Bryan, J. D., & Rosse, R. L. (1983). Validity of Army recruiter behavioral assessment: Does the assessor make a difference? Journal of Applied Psychology, 63(3).

Borman, W. C., Hough, L. M., & Dunnette, M. D. (1976). Development of behaviorally based rating scales for evaluating the performance of U.S. Navy recruiters (NPRDC Technical Report 76-31). San Diego: Navy Personnel Research and Development Center.

Borman, W. C., Rosse, R. L., & Abrahams, N. M. (1980). An empirical construct validity approach to studying predictor-job performance links. Journal of Applied Psychology, 65, 662-671.

Borman, W. C., Rosse, R. L., & Rose, S. R. (1983). Development and validation of an inventory battery to predict performance in Navy officer recruiting. Minneapolis, MN: Personnel Decisions Research Institute.

Borman, W. C., Rosse, R. L., & Toquam, J. L. (1982). The impact of environmental factors and considerations of recruit quality on Navy recruiter production. Minneapolis: Personnel Decisions Research Institute.

Borman, W. C., Russell, T. L., & Skilling, N. J. (1986). Development of behavior-based rating scales and analysis of recruiter selection battery data for the Army recruiter job. Report submitted to the U.S. Army Research Institute, Alexandria, VA.

Borman, W. C., Toquam, J. L., & Rosse, R. L. (1976b). Dimensions of the Army recruiter and guidance counselor job. Minneapolis, MN: Personnel Decisions Research Institute.

Borman, W. C., Toquam, J. L., & Rosse, R. L. (1979). An inventory battery to predict Navy and Marine Corps recruiter performance: Development and validation. Minneapolis: Personnel Decisions Research Institute.

Borman, W. C., White, L. A., Gast, I. F., & Pulakos, E. D. (August, 1985). Performance ratings as criteria: What is being measured? Paper presented at the annual meeting of the American Psychological Association, Los Angeles, CA.

Bownas, D. A., Rosse, R. L., & Dunnette, M. D. (1977). Construct validation of a selection battery for the entry-level firefighter position. Minneapolis, MN: Personnel Decisions Research Institute.

Bray, D. W., & Campbell, R. J. (1968). Selection of salesmen by means of an assessment center. Journal of Applied Psychology, 52, 36-41.

Brogden, H. E., & Taylor, E. K. (1949). The validity of recruiter selection instruments at various points of cut (Report No. 781). Washington, DC: Personnel Research Section.

Brown, G. H., Wood, M. D., & Harris, J. D. (1975). Army recruiters: Criterion development and preliminary validation of selection procedure (FR-ED-75-8). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Campbell, J. P., Dunnette, M. D., Lawler, E. E., III, & Weick, K. E., Jr. (1970). Managerial behavior, performance, and effectiveness. New York: McGraw-Hill.

Churchill, G. A., Jr., Ford, N. M., Hartley, S. W., & Walker, O. C., Jr. (1985). The determinants of salesperson performance: A meta-analysis. Journal of Marketing Research, 22, 103-118.

Churchill, G. A., Jr., Ford, N. M., & Walker, O. C., Jr. (1981). Sales force management: Planning, implementation and control. Homewood, IL: Richard D. Irwin, 272-296.

Cocanougher, A. B., & Ivancevich, J. M. (1978). "BARS" performance rating for sales force personnel. Journal of Marketing, 42, 87-95.

Cleveland, E. A. (1948). Sales personnel research: 1935-1945: A review. Personnel Psychology, 1, 211-255.

Cotham, J. C., III. (1969a). Predicting salesmen's performance by multiple discriminant analysis. Southern Journal of Business, 4, 25-34.

Cotham, J. C., III. (1969b). Using personal history information in retail salesman selection. Journal of Retailing, 45, 31-38, 84.

Cravens, D. W., & Woodruff, R. B. (1973). An approach for determining criteria of sales performance. Journal of Applied Psychology, 57, 242-247.

Cravens, D. W., Woodruff, R. B., & Stamper, J. C. (1972). An analytical approach for evaluating sales territory performance. Journal of Marketing, 36, 31-37.

Cunningham, R. M. (1935). Some problems in measuring performance of industrial salesmen. Harvard Business Review, 14, 98-112.

Dubinsky, A. J., & Hartley, S. W. (in press). Antecedents of retail salesperson performance: A path-analytic perspective. Journal of Business Research.

Dugan, R. D. (1960). Evaluating territorial sales efforts. Journal of Applied Psychology, 44, 107-110.

Dunnette, M. D., & Kirchner, W. K. (1959). A job description check list for differentiating different kinds of sales jobs. Personnel Psychology, 12, 421-429.

Dunnette, M. D., & Kirchner, W. K. (1960). Psychological test differences between industrial salesmen and retail salesmen. Journal of Applied Psychology, 44, 121-125.

Dunnette, M. D., Kirchner, W. K., & DeGidio, J. (1958). Relations among scores on the Edwards Personal Preference Schedule, the California Psychological Inventory, and the Strong Vocational Interest Blank for an industrial sample. Journal of Applied Psychology, 42, 178-181.

Dunnette, M. D., McCartney, J., Carlson, H. C., & Kirchner, W. K. (1962). A study of faking behavior on a forced-choice self-description checklist. Personnel Psychology, 15, 13-24.

Ferguson, L. W. (1960). Ability, interest, and aptitude. Journal of Applied Psychology, 44, 126-131.

Flanagan, J. C. (1954). The critical incident technique. Psychological Bulletin, 51, 327-358.

Gable, M., Mattheiss, T. H., & Muczyk, J. P. (1984). Predicting the success of salesmen through the use of a forced choice personality test and discriminant analysis. Akron Business and Economic Review, 5, 30-34.

Ghiselli, E. E. (1969). Prediction of success of stockbrokers. Personnel Psychology, 22, 125-130.

Graf, R. G., & Brower, D. B. (1976). The development of an interest inventory for the selection of Marine Corps recruiters. San Diego: Navy Personnel Research and Development Center.

Greenberg, H., & Mayer, D. (1964). A new approach to the scientific selection of successful salesmen. The Journal of Psychology, 57, 113-123.

Guion, R. M., & Gottier, R. F. (1965). Validity of personality measures in personnel selection. Personnel Psychology, 18, 135-164.

Harrell, T. W. (1960). The validity of biographical data items for food company salesmen. Journal of Applied Psychology, 44, 31-33.

Harrell, T. W. (1960b). The relation of test scores to sales criteria. Personnel Psychology, 44, 65-69.

Harris, W. A., & Vincent, N. L. (1967). Comparison of performance of sales training graduates and nongraduates. Journal of Applied Psychology, 51, 436-441.

Hughes, J. L. (1956). Expressed personality needs as predictors of sales success. Personnel Psychology, 9, 347-357.

Husband, R. W. (1949). Techniques of salesmen selection. Educational and Psychological Measurement, 9, 120-148.

Kanfer, R., & Borman, W. C. (1986). Predicting salesperson performance: A review of the literature. Personnel Decisions Research Institute to the U.S. Army Research Institute for the Behavioral and Social Sciences.

Katzell, R., Barrett, R. S., & Parker, T. C. (1961). Job satisfaction, job performance, and situational characteristics. Journal of Applied Psychology, 45, 65-72.

Kennedy, J. E. (1958). A general device versus more specific devices for selecting car salesmen. Journal of Applied Psychology, 42, 206-209.

Kerber, K. W., Campbell, James P., & Lapide, L. (1985). Time use, performance, and organizational turnover among computer sales representatives. Paper presented at APA Meetings, Los Angeles, CA.

Kirchner, W., & Dunnette, M. D. (1957). Identifying the critical factors in successful salesmanship. Personnel, 34, 54-59.

Krug, S. E. (1972). Psychological and demographic predictors of success as a Naval recruiter (final report). Washington, DC: Navy Recruiting Command.

Lamont, L. M., & Lundstrom, W. J. (1977). Identifying successful industrial salesmen by personality and personal characteristics. Journal of Marketing Research, 14, 517-529.

Landy, F. J. & Farr, J. L. (1983). The measurement of work performance. New York: Academic Press.

Larriva, R. F. (1975). U.S. Marine Corps recruiter performance prediction study. Unpublished manuscript.

Massey, I. H., & Mullins, C. J. (1966). Validation of the recruiter-salesman selection test. Lackland Air Force Base, TX: Personnel Research Laboratory, Aerospace Medical Division, Air Force Systems Command.

Matteson, M. T., Ivancevich, J. M., & Smith, S. V. (1984). Relation of Type A behavior to performance and satisfaction among sales personnel. Journal of Vocational Behavior, 25, 203-214.

Mayfield, E. C. (1972). Value of peer nominations in predicting life insurance sales performance. Journal of Applied Psychology, 56, 319-323.

Merenda, P. F., & Clarke, W. V. (1959). The predictive efficiency of temperament characteristics and personal history variables in determining success of life insurance agents. Journal of Applied Psychology, 43, 360-366.

Merenda, P. F., Clarke, W. V., & Hall, C. E. (1961). Cross-validity of procedures for selecting life insurance salesmen. Journal of Applied Psychology, 45, 376-380.

Miner, J. B. (1962). Personality and ability factors in sales performance. Journal of Applied Psychology, 46, 6-13.

Mosel, J. N. (1952). Prediction of department store sales performance from personal data. Journal of Applied Psychology, 36, 8-10.

Mosel, J. N., & Wade, R. R. (1951). A weighted application blank for reduction of turnover in department stores sales clerks. Personnel Psychology, 4, 177-184.

Murray, L. E., & Bruce, M. M. (1959). A study of the validity of the sales comprehension test and sales motivation inventory in differentiating high and low production in life insurance selling. Journal of Applied Psychology, 43, 246-248.

Pace, R. W. (1962). Oral communication and sales effectiveness. Journal of Applied Psychology, 46, 321-324.

Peterson, N. G., & Houston, J. S. (1980). The prediction of correctional officer job performance: Construct validation in an employment setting. Minneapolis, MN: Personnel Decisions Research Institute.

Pulakos, E. D., & Borman, W. C. (Ed.). (1985). Development and field test of Army-wide rating scales and the rater orientation and training program. (Report #716). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Rodgers, D. A. (1959). Personality of the route salesman in a basic food industry. Journal of Applied Psychology, 43, 235-239.

Ruch, F. L., & Ruch, W. W. (1967). The K factor as a (validity) suppressor variable in predicting success in selling. Journal of Applied Psychology, 51, 201-204.

Rush, C. H., Jr. (1953). A factorial study of sales criteria. Personnel Psychology, 6, 9-24.

Smith, P. C., & Kendall, C. M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. Journal of Applied Psychology, 47, 149-155.

Tanofsky, R., Shepps, R. R., & O'Neill, P. J. (1969). Pattern analysis of biographical predictors of success as an insurance salesman. Journal of Applied Psychology, 53, 136-139.

Walker, O. C., Jr., Churchill, G. A., Jr., & Ford, N. M. (1977). Motivation and performance in industrial selling: Present knowledge and needed research. Journal of Marketing Research, 14, 156-168.

Walker, O. C., Churchill, G. A., Jr., & Ford, N. M. (1979). Where do we go from here? Selected conceptual and empirical issues concerning the motivation and performance of the industrial salesforce. In G. Albaum & G. A. Churchill, Jr. (Eds.), Critical issues in sales management: State-of-the-art and future research needs. Eugene, OR: College of Business Administration, University of Oregon, pp. 10-75.

Waters, L. K., & Waters, C. W. (1970). Peer nominations as predictors of short-term sales performance. Journal of Applied Psychology, 54, 42-44.

Weaver, C. N. (1969). An empirical study to aid in the selection of retail salesclerks. Journal of Retailing, 45, 22-26.

Weitz, B. A. (1979). A critical review of personal selling research: The need for contingency approaches. In G. Albaum & G. A. Churchill, Jr. (Eds.), Critical issues in sales management: State-of-the-art and future research needs. Eugene, OR: College of Business Administration, University of Oregon, pp. 76-126.

Wollack, L., & Kipnis, D. (1960). Development of a device for selecting recruiters (Task assignment PF-016-05-003-W2). Washington, DC: U.S. Naval Personnel Research Field Activity.

Zdep, S. M., & Weaver, H. B. (1967). The graphoanalytic approach to selecting life insurance salesmen. Journal of Applied Psychology, 51, 295-299.

APPENDIX A

Summary of the Development of the  
Special Assignment Battery (SAB)

## Appendix A

### Summary of the Development of the Special Assignment Battery (SAB)

Over the years, extensive research has been conducted to delineate the personal attributes related to recruiter success. A number of these studies have employed a specific battery of vocational interest, biographical, and personality items called the Special Assignment Battery (SAB). Because the SAB has proven successful as a predictor of recruiter performance in several research efforts and appears to be a good starting point for future recruiter selection research, this appendix summarizes its development and subsequent use in research.

The development of the SAB began with a job analysis wherein researchers learned about the recruiter job and developed hypotheses about the personal characteristics (e.g., personality, interests, background) related to recruiter performance (Borman, Hough & Dunnette, 1976). Through numerous interviews and workshops with recruiters, the research team also developed behavior-based performance appraisal rating scales for recruiters. Ratings of recruiter performance on these rating scales have served as the criteria for evaluating the validity of the SAB in a number of studies and were instrumental in the development of the SAB.<sup>1</sup> We, therefore, discuss the development of the rating scales briefly in the following paragraphs and then summarize the development of the SAB itself.

#### Development of Behavior-Based Recruiter Rating Scales

To develop behavior-based rating scales for recruiters we used the behavior analysis, or critical incident methodology. Over 800 examples of effective and ineffective recruiter performance were gathered from field recruiters, their supervisors, and recruits. Performance categories were then developed on the basis of the content of the incidents, and recruiters and their supervisors categorized the incidents and rated the effectiveness each represented. Finally, Behavior Summary Scales (BSS) were developed by grouping reliably sorted and rated incidents into each dimension and writing behavior summary statements reflecting the content of those incidents.

The resulting Navy recruiter scales consisted of eight dimensions or categories of performance. They are:

1. Locating and Contacting Prospects,
2. Gaining and Maintaining Rapport,
3. Obtaining Information from Prospects and Making Good Person-Navy Fits,

---

<sup>1</sup>In 1985, the Army administered the Recruiter Selection Battery-Experimental (RSB-X) to over 400 recruiters. This battery, the RSB-X, is a slight modification of the SAB. A few of the biographical items were tailored to the Army, and the interest inventory was omitted.

4. Salesmanship Skills,
5. Establishing and Maintaining Good Relationships in the Community,
6. Providing Knowledgeable and Accurate Information about the Navy,
7. Administrative Skills, and
8. Supporting Other Recruiters and the Command.

In order to evaluate these rating scales, supervisory, peer, and self-performance ratings were collected and analyzed for 267 Navy and 118 Marine Corps recruiters (Borman et al., 1979). Interrater reliabilities obtained for the pooled peer-supervisor-self-ratings were  $r = .57$  for the Navy and  $r = .48$  for the Marine Corps scales. Analysis of rating errors (i.e., leniency, restriction-of-range, and halo) showed results comparable to those discussed previously concerning the summary scale format. When the pooled ratings were factor analyzed, a three-factor solution resulted as described below:

1. Selling Skills - selling Navy effectively to prospects; displaying confidence and effectiveness in the recruiting sequence - prospecting, selling, and closing;
2. Human Relations Skills - establishing and maintaining good interpersonal relations with prospects, recruits, and persons in the community; and
3. Organizing Skills - planning ahead and organizing time efficiently, completing paperwork accurately and on time.

Subsequent research with these scales (Borman et al., 1981) replicated the three-factor solution shown above. Therefore, the scales appear to be reliable, robust to various rating errors, and representative of a stable and psychologically meaningful factor structure.

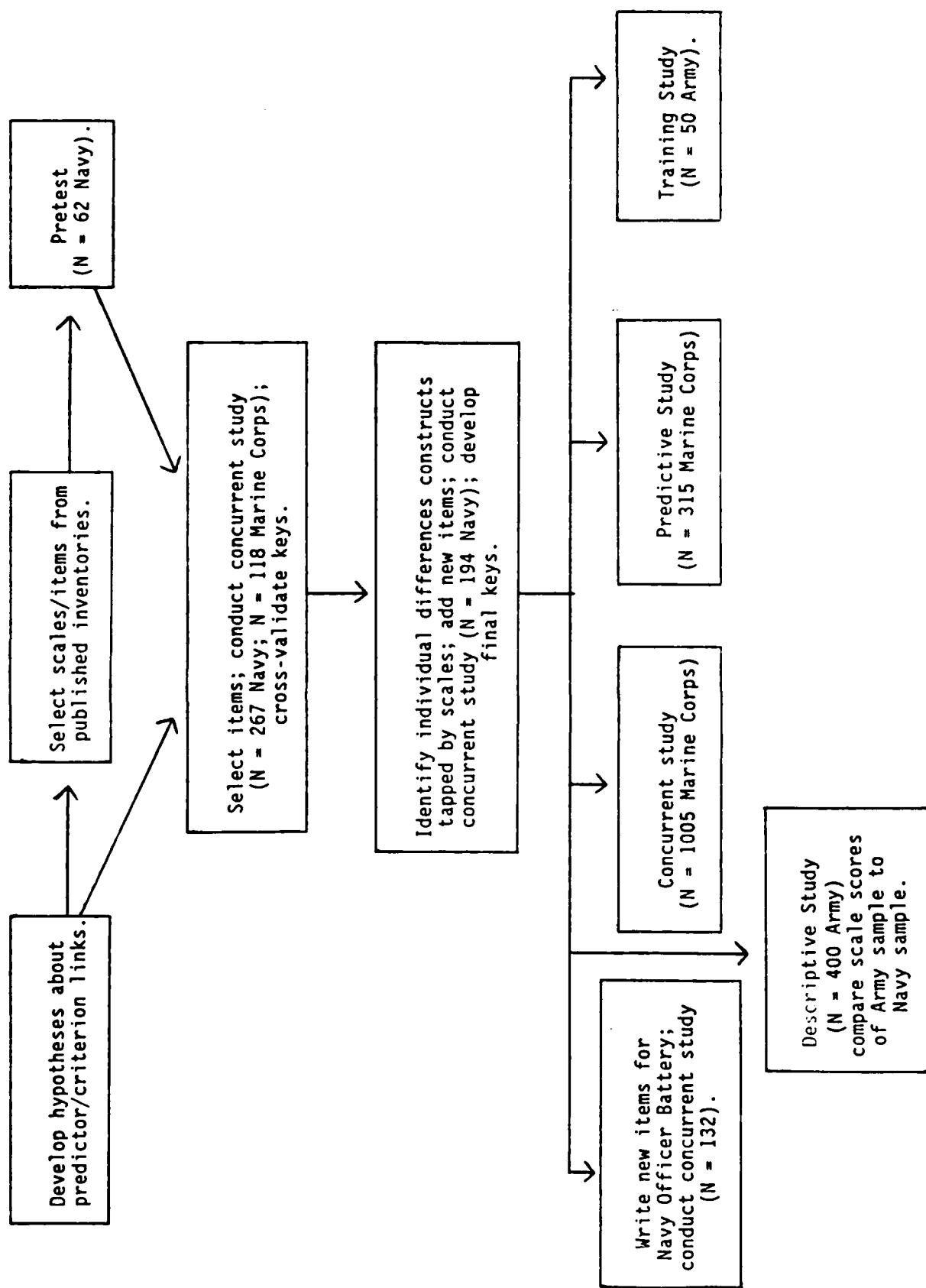
#### Development of the Special Assignment Battery

Hypothesized (and later, empirical) relationships between the performance factors described above and personality, interest and background items were used to formulate the SAB and its keys. As shown in Figure A-1 hypothesized links were first used to select items and scales from published inventories, and item-performance correlations from pretest data were used to further refine the item pool. At this point, four keys were developed for the SAB (i.e., one for each performance category that the items were expected to predict).<sup>2</sup> A concurrent validation study was then

---

<sup>2</sup>The SAB keys are named after the performance category they are intended to predict. Validity estimates are correlations between the performance construct (e.g., Selling Skills) and the SAB key (e.g., Selling Skills) designed to predict that aspect of performance.

Figure A-1. Flow Diagram of Major Steps in the Development of and Research with the Special Assignment Battery (SAB)



conducted, and the keys were cross-validated via a Monte Carlo procedure (Borman, Toquam, and Rosse, 1979). Next, the items within each key were factor analyzed to identify the individual differences constructs tapped by each key, and new items, targeted toward tapping the same non-cognitive constructs were written. In turn, the new items along with the old items were validated in a concurrent study (Borman, Rosse, and Abrahams, 1980), and final keys were developed.

### Research Employing the SAB

Subsequent research employing the SAB can be categorized into two types: (1) developmental/refinement studies tailoring or examining SAB items, and (2) attempts to validate SAB scores against various criteria.

Developmental research. In one study (Borman, Rosse, & Rose, 1983) the SAB items were refined and tailored to the Navy officer recruiter job. Likewise, one goal of the Army's ongoing research is to examine Army recruiters' SAB scores (RSB-X) and compare them to those of Navy recruiters.

Validation research. Recruiter performance, production, and training success have served as criteria for validating the SAB. Table A-1 presents the validity coefficients obtained in these studies. In sum, the SAB has consistently predicted recruiter performance (i.e., self-supervisor-peer ratings) across studies.

Attempts to validate the SAB against a production criterion have also been successful. In one study the SAB validated poorly against production (Borman et al., 1979). This finding is, however, inconsistent with later research (Atwater, Abrahams, & Trent, 1986; Borman, Rosse, & Abrahams, 1980).

In two studies the SAB has not validly predicted recruiter success in training. Thus, it appears that the individual differences constructs that lead to, or predict, recruiter performance and production on-the-job differ somewhat from those characterizing recruiters who are successful in training. Indeed, the behaviors that make for success in training (e.g., studying) may be quite different from those leading to job success. It is likely that cognitive variables, as well as non-cognitive, are important for training performance, and the validation of assessment center exercises tapping both cognitive and non-cognitive constructs against training criteria (Borman, 1982; Borman, Eaton & Rosse, 1983) supports this view.

Table A-1. Summary of SAB Validity Estimates Across Studies

## Ratings Criteria

## SAB Scales/Performance Category

Study	Sample	Item Type	Selling Skills	Human Relations	Organizing Skills	Overall Performance	Recpot	Composite Across One Type of Item
1.	Navy 267 MC 118	Personality	21 19	16 19	29 38	21 24	--	--
	Navy 267 MC 118	Interest	19 19	17 24	26 20	17 26	--	--
	Navy 267 MC 118	Biographical	16 -01	05 10	23 15	23 15	--	--
2.	Navy 194	Old plus New Personality Old plus New Interest Final Keys Personality & Interest Personality	23 21 43 15	24 22 46 30	15 12 40 18	22 22 43 19	--	--
3.	Navy 132 Officer Recruiter	Interest Final Keys Personality & Interest	16 19	11 15	07 13	16 20	--	--
5.	MC 1005	Personality					20	--
Production Criteria								
1.	Navy 267 MC 118	Personality	--	--	--	--	--	01
	Navy 267 MC 118	Interest	--	--	--	--	--	24
	Navy 267 MC 118	Biographical	--	--	--	--	--	-02
2.	Navy 194	Final Keys	22	23	13	26	27	--
5.	MC 1005 MC 315	Final Keys	23 23	15 15	09 09	24 24	22 22	--
Training Criteria/Final Keys								
4.	Army 50	Criterion	02	-06	-01	-09	--	--
	Army 46	Phase 1 Performance	17	15	07	15	--	--
	Army 46	Phase 2 Performance	-12	-27	-15	-09	--	--
5.	MC 431	Time to Complete Completion/Failure	--	--	--	--	08	--

Studies

1. Cross-validities obtained via Monte Carlo procedure by Borman, Toquam, and Rosse (1979).
2. Borman, Rosse, and Abrahams (1980); final key validities were not cross-validated.
3. Borman, Rosse, and Rose (1983).
4. Borman (1982).
5. Atwater, Abrahams, and Trent (1986).